2

Control System Analysis and Synthesis Based on Matrix Inequalities

Dissertation

Submitted in partial fulfillment of the requirements for the degree of Doctor of Informatics

in

Department of Applied Analysis and Complex Dynamical Systems

by

Yuji Wakasa

Kyoto University

June 2000

Abstract

Linear (or bilinear) matrix inequalities are inequalities of the form: an affine (or a biaffine) combination of symmetric matrices is positive semidefinite. Since the advent of the Lyapunov stability theory, it has been recognized that a number of control problems can be described by such matrix inequalities, but they had remained numerically intractable for a long time. However, due to the recent progress of performance of computers and development of efficient algorithms such as an interior-point algorithm, linear matrix inequalities have become numerically solvable in the last decade. Under this situation, various ways to formulate control problems as linear matrix inequalities have been intensively studied. This is the so-called linear matrix inequality approach. Moreover, bilinear matrix inequalities which are more general than linear matrix inequalities have received attention because of the ability to naturally describe various control synthesis problems. In this thesis, we investigate control system analysis and synthesis by linear matrix inequalities and a numerical solution to bilinear matrix inequalities.

We first present an analysis method for estimating sensitivity of performance of control systems with uncertainty. Here noticing that many control performance analysis problems are reduced to semidefinite programming problems, i.e., a problem that minimizes a linear function subject to linear matrix inequality constraints, we take two steps for the goal. In the first step, we propose a new form of complementarity condition and derive results of sensitivity analysis of semidefinite programming. In the second step, we apply the obtained results to sensitivity analysis of control systems.

We then propose two types of control system synthesis methods based on linear matrix inequalities. One is concerned with control system synthesis considering a tradeoff between evaluated uncertainty ranges and control performance. We here reduce the syn-

thesis problem to a semi-infinite programming problem with bilinear matrix inequality constraints. An approximate method for the problem is presented, and its convergence is proved. The other is concerned with robust model predictive control with rate constraints. We here present linear matrix inequality conditions for rate constraints of inputs and outputs in the framework of a recently proposed model predictive control method based on linear matrix inequalities. By a numerical example, we show that a good performance is obtained in practice by using the presented linear matrix inequality conditions.

As a solution to bilinear matrix inequalities, we present a global optimization algorithm based on the primal-relaxed dual method that is a global optimization method. We also modify the algorithm from the viewpoint of computational efficiency. A numerical example is given to illustrate the geometrical interpretation and effectiveness of the proposed method.

Acknowledgments

I would like to express my sincere gratitude to all those who helped me and contributed in various ways toward completion of this thesis.

First of all, I would like to thank my current supervisor, Professor Yutaka Yamamoto for giving me his very detailed and valuable comments on the earlier drafts of this thesis. He also has been teaching me from how to write academic papers to what is researcher's philosophy since I joined in his laboratory about two years ago. It is a great pleasure to me in my research life.

I would also like to express my special thanks to my previous supervisor, Professor Tohru Katayama for having given me his continual guidance and encouragement since I was a senior student. Thanks to him, I had come to take interest in research on system and control theory.

I am very grateful to Professor Tetsuzo Tanino at Osaka University for his valuable advice and suggestions on optimization. I could study fundamentals of optimization theory and methods while I worked at his laboratory. Most of this thesis are based on works of those days.

My gratitude should be extended to Associate Professor Hisaya Fujioka and all members of the intelligent and control systems laboratory for their academic stimulation and daily friendship.

Special thanks are addressed toward my parents for their emotional support and understanding to my study. Finally, I bless my wife for her devotion, encouragement and affection.

Contents

Notation

$\Re$	real numbers
$\Re^n$	n-dimensional real vectors
$\Re^{m\times n}$	$m \times n$ real matrices
$\lambda_{\max}(A)$	maximum eigenvalue of a symmetric matrix A
$\lambda_{\min}(A)$	minimum eigenvalue of a symmetric matrix A
$\sigma_{\max}(A)$	maximum singular value of a matrix A
$\sigma_{\min}(A)$	minimum singular value of a matrix A
$\mathbf{Tr}\,A$	trace of a matrix A
$\mathrm{rank}(A)$	rank of a matrix A
$\mathrm{diag}(A_1,\ldots,A_n)$	block-diagonal matrix with A_i as its i-th block-diagonal element
$[a_{ij}]$	matrix with a_{ij} as its i-th row and j-th column element
$A^\dagger$	Moore-Penrose inverse of A
$\mathbf{Co}\{A_1,\ldots,A_n\}$	convex hull of matrices $A_1,\ldots,A_n$
I_n	$n \times n$ identity matrix
$O_{m\times n}$	$m \times n$ zero matrix
0_n	$n \times n$ zero matrix

To reduce the number of parentheses required, we adopt the convention that operators **Tr** has lower precedence than multiplication, transpose, etc. Thus $\mathbf{Tr}\,A^TB$ means $\mathbf{Tr}(A^TB)$.

Acronyms

ARE	algebraic Riccati equation
BMI	bilinear matrix inequality
LMI	linear matrix inequality
LQG	linear quadratic Gaussian
LQR	linear quadratic regulator
MPC	model predictive control
SDP	semidefinite programming, semidefinite program

Chapter 1

Introduction

This thesis is concerned with the control system analysis and synthesis based on linear matrix inequalities (LMIs) and bilinear matrix inequalities (BMIs). In the past decade, LMIs have received an increasingly broader acceptance as a useful tool for control system analysis and synthesis. On the other hand, BMIs has been intensively investigated as a more general framework of LMIs these days. In this chapter, we first review a general historical background of such matrix inequalities approach and then summarize the contributions and organization of this thesis.

1.1 Background of LMIs and BMIs in Control Theory

The history of LMIs and BMIs in the analysis of dynamical systems goes back more than a century. The research began in about 1890, when Lyapunov published his seminal work introducing what we now call the Lyapunov theory. He showed that the differential equation

$$\frac{d}{dt}x(t) = Ax(t) \tag{1.1}$$

is stable, i.e., all trajectories converge to zero, if and only if there exists a positive definite matrix P such that

$$A^T P + PA < 0. \tag{1.2}$$

The requirement $A^TP+PA<0$ (with $P>0$) is what we now call the *Lyapunov inequality* on P, which is a special form of an LMI. He also showed that this first LMI could be explicitly solved. Indeed, we can pick any $Q=Q^T>0$ and then solve the linear equation $A^TP+PA=-Q$ for a matrix P, which is guaranteed to be positive definite if system (1.1) is stable. In summary, the first LMI used to analyze the stability of a dynamical system was the Lyapunov inequality (1.2), which can be solved analytically by solving a set of linear equations.

Since then many different kinds of LMIs have been introduced in control theory, but only those of relatively small order were solvable by hand before 1950's. The next major breakthrough came in 1970's, when specific families of LMIs such as the LMIs appearing in the so-called positive real lemma were shown to be solvable, regardless of the size, by solving a certain algebraic Riccati equation (ARE). In a 1971 paper [64] on the quadratic optimal control, J. C. Willems was led to the LMI

$$\begin{bmatrix} A^TP+PA+Q & PB+C^T \\ B^TP+C & R \end{bmatrix} \geq 0,$$

and pointed out that it can be solved by studying the symmetric solutions of the ARE

$$A^TP+PA-(PB+C^T)R^{-1}(B^TP+C)+Q=0,$$

which in turn can be found by exhibiting an eigenstructure of a related Hamiltonian matrix. This type of methods is a "closed-form" or "analytic" solution that can be used to solve special forms of LMIs, since the standard algorithms that solve it are very predictable in terms of the effort required, which depends almost entirely on the problem size and not the particular problem data.

In the early 1980's, it was recognized that the LMIs that arise in system and control theory can be formulated as convex optimization problems that are amenable to computer calculation. Although this is a simple observation, it had some important consequences, the most important of which was that we could reliably solve many LMIs for which no analytic solution had been found (or was likely to be found).

In 1984, Karmarkar [32] introduced a new linear programming algorithm that solves linear programs in polynomial-time, like the ellipsoid method, but in contrast to the ellipsoid method, is also very efficient in practice. Karmarkar's work spurred an enormous

amount of work in the area of interior-point methods for linear programming (including rediscoveries of efficient methods that were developed but ignored in the 1960's). Essentially, all of this research activity were concentrated on algorithms for linear and convex quadratic programs.

Then in 1988, Nesterov and Nemirovskii [41] showed that the interior-point methods for linear programming can, in principle, be generalized to all convex optimization problems. The key element is whether there exits a barrier function with a certain property called self-concordance [42, 39]. LMIs are an important class of convex constraints for which readily computable self-concordant barrier functions are known, and therefore interior-point methods are applicable.

Independently of Nesterov and Nemirovskii, several researchers generalized interior-point methods from linear programming to the so-called semidefinite programming (SDP)[1]. Special classes of the SDP have a long history in optimization. For example, certain eigenvalue minimization problems that can be cast into the frame of SDPs have been studied in combinatorial optimization [45]. The efficiency of recent interior-point methods for SDP, which is directly responsible for the popularity of SDP in control, has therefore also attracted a great deal of interest in optimization. Vast progress has been made in the last five years, and today almost all interior-point methods for linear programming have been extended to the SDP [54].

The SDP is a problem of the form

$$\min_x \{c^Tx \mid F_0+\sum_{i=1}^m x_iF_i \geq 0\}. \tag{1.3}$$

The problem data are the vector $c\in\Re^m$ and $m+1$ symmetric matrices $F_0,\ldots,F_m\in\Re^{n\times n}$. The inequality sign in (1.3) means that the symmetric matrix is positive semidefinite. Though the form of the SDP (1.3) appears very specialized, it turns out that it is frequently encountered in systems and control. Examples include multicriterion LQG/LQR, synthesis of linear state-feedback for multiple plants, robustness analysis and robust controller design, gain-scheduling, and many others [8].

[1] We shall use SDP to mean *semidefinite programming* as well as a *semidefinite program*, i.e., a *semidefinite programming problem*.

While the LMI approach is successful in the field of control, there still exist many important control problems which cannot (or are not known to) be equivalently expressed as SDPs. For example, the low order controller synthesis [26], multi-objective control and structure [48], distributed control synthesis [29], simultaneous optimization of control and structure [43]. Even if these problems can be reduced to SDPs by some relaxation methods, the resultant control performance is often conservative. However, many of such control problems are naturally expressed as BMIs. In view of this observation, Safonov [47] and Goh [24] introduced BMIs as a unified description of a wide variety of control problems. Since then, several researchers have tried to develop algorithms for solving BMIs [23, 16, 49, 33, 62, 4]. The BMI is the following form.

$$F_{00}+\sum_{i=1}^{n_x} x_i F_{i0}+\sum_{j=1}^{n_y} y_j F_{0j}+\sum_{i=1}^{n_x}\sum_{j=1}^{n_y} x_i y_j F_{ij} \geq 0, \tag{1.4}$$

where $x \in \Re^{n_x}$, $y \in \Re^{n_y}$ are variable vectors and coefficient matrices $F_{ij} \in \Re^{m\times m}$ are symmetric. For fixed y the BMI (1.4) reduces to an LMI in the variable x; for fixed x it reduces to an LMI in the variable y. The fundamental difference from LMIs is that the BMI problem is nonconvex, and no non-exponential time algorithms for its solution are known to exist. In fact, it has been shown that the BMI problem is NP-hard [51]. Nevertheless, some numerical algorithms for BMIs have been investigated intensively with the aim of solving control design problems of practical size. Existing BMI methods are either local methods that alternate between optimization over y and over x [23, 21], or global methods such as branch and bound methods [23, 16, 49, 18], primal-relaxed dual methods [62, 4], and d.c. programming [36].

1.2 Contributions and Organization of This Thesis

The results obtained in this thesis are summarized as follows.

- We propose some new methods for control system analysis and synthesis via LMIs.

  In this thesis, we deal with LMI approaches to control problems from two viewpoints. One is an application of analysis methods in optimization theory. We propose a

  method for estimating sensitivity of SDP with uncertainty and apply the method to sensitivity analysis of control systems. The other is to formulate new control objectives as LMIs. As described above, development of numerical optimization involving LMIs not only gave a unified approach to conventional control problems but also made it possible to solve new control problems. We propose a control design method considering a tradeoff in robust control and present LMI conditions for rate constraints in the framework of state-feedback control.

- A global optimization method for solving BMIs is developed.

  While LMIs have become powerful tools for many control problems, there still exist some important control problems which cannot (or are not likely to) be described by LMIs. For such control problems, BMIs are more general and natural frameworks. However, it has been recognized to be difficult to efficiently solve BMIs due to their biconvexity. In this thesis, we present a global optimization algorithm for the BMI problem based on the primal-relaxed dual method. This method is a global optimization method for mathematical programming problems whose objective function and constraints are both biconvex.

This thesis consists of two parts. In Chapters 3–5, we investigate control system analysis and synthesis based on LMIs. Chapter 6 are concerned with methods for solving BMIs. The organization of this thesis is as follows.

Chapter 2 gives preliminaries required in the following chapters. We describe the Schur complement, the $\mathcal{S}$-procedure and elimination lemmas, which are useful in formulating LMIs for control problems. We also present some properties of LMIs and BMIs such as convexity and biconvexity, and give simple examples of LMIs and BMIs in control theory.

Chapter 3 is concerned with sensitivity analysis of control systems with uncertainty. To this end, we first study the effect of perturbations in the SDP on the optimal solution and the optimal value function. A new form of complementarity conditions is proposed and the first-order partial derivatives of the optimal value function with respect to parametric variation are explicitly expressed by the problem data and the optimal solution.

Furthermore, the above result is applied to the sensitivity analysis of control systems with parametric uncertainties.

Chapter 4 and 5 deal with control system synthesis based on LMIs.

In **Chapter 4**, we presents a design method of control systems such that a designer can flexibly take account of tradeoffs between evaluated uncertainty ranges and the level of control performance. The problem is reduced to a BMI problem and approximately solved by a sequence of LMIs, and the convergence of the proposed approximation is proved. A numerical example shows the effectiveness of the proposed method in comparison with the standard robust control.

Recently, an LMI-based model predictive control (MPC) technique has been proposed [35]. Many MPC techniques usually take into account constraints such as amplitude and rate limits and have therefore been very successful in industry. However, the MPC method [35] deals with only amplitude limits but not rate limits. In **Chapter 5**, we present LMI conditions for the rate limits in the framework of the LMI-based MPC technique. These additional constraints make the LMI-based MPC technique more practical since the rate limits, as well as the amplitude limits, can be taken into consideration.

In **Chapter 6**, we present a global optimization algorithm for the BMI problem based on the primal-relaxed dual method. We also modify the algorithm from the viewpoint of computational efficiency. Numerical examples are given to illustrate the geometrical interpretation and effectiveness of the proposed method.

In **Chapter 7**, we state the conclusion of this thesis and discuss future directions of our research.

Chapter 2

Preliminaries

This chapter gives preliminaries required in the following chapters. We provide some lemmas for matrix inequalities and show simple examples of LMIs and BMIs.

2.1 Some Lemmas for Matrix Inequalities

A variety of control problems can be formulated as LMIs and BMIs by using techniques such as Schur complements, the $\mathcal{S}$-procedure and elimination lemmas. Here we list some useful lemmas for matrix inequality formulation.

2.1.1 Schur Complements

The so-called Schur complements [8, 30] are useful in converting nonlinear (convex) inequalities to LMI form. The case of strict inequalities is as follows.

Lemma 2.1 *The following conditions are equivalent for a real symmetric matrix Θ:*

(i)

$$\Theta := \begin{bmatrix} \Theta_{11} & \Theta_{12} \\ \Theta_{12}^T & \Theta_{22} \end{bmatrix} > 0.$$

(ii)

$$\Theta_{11} > 0 \ \textit{and} \ \Theta_{22} - \Theta_{12}^T \Theta_{11}^{-1} \Theta_{12} > 0.$$

(iii)

$$\Theta_{22} > 0 \text{ and } \Theta_{11} - \Theta_{12}\Theta_{22}^{-1}\Theta_{12}^T > 0.$$

The above result can be generalized to nonstrict inequalities as follows [8].

Lemma 2.2 *Suppose Q and R are symmetric. The condition*

$$\begin{bmatrix} Q & S \\ S^T & R \end{bmatrix} \geq 0$$

is equivalent to

$$R \geq 0, \quad Q - SR^\dagger S^T \geq 0, \quad S(I - RR^\dagger) = 0,$$

where $R^\dagger$ denotes the Moore-Penrose inverse of R.

2.1.2 The $\mathcal{S}$-Procedure

We often encounter a constraint where some quadratic function (or quadratic form) be negative whenever some other quadratic functions (or quadratic forms) are all negative. In some cases, this constraint can be expressed as an LMI in the data defining quadratic functions or forms; in other cases, we can form an LMI that is a conservative but often useful approximation of the constraint. The $\mathcal{S}$-procedure is a technique for expressing such a constraint as an LMI.

Let $T_0, \ldots, T_p \in \Re^{n\times n}$ be symmetric matrices. Consider the following condition on $T_0, \ldots, T_p$:

$$\zeta^T T_0 \zeta < 0 \text{ for all } \zeta \neq 0 \text{ such that } \zeta^T T_i \zeta \leq 0, \quad i = 1, \ldots, p. \tag{2.1}$$

It is obvious that if

$$\text{there exists } \tau_1 > 0, \ldots, \tau_p > 0 \text{ such that } T_0 - \sum_{i=1}^{p} \tau_i T_i < 0, \tag{2.2}$$

then (2.1) holds. Note that (2.2) is an LMI in the variables T_0 and $\tau_1, \ldots, \tau_p$. It is a nontrivial fact that when $p = 1$, the converse holds, however, the following lemma holds [8].

Lemma 2.3 *Let $T, S \in \Re^{n\times n}$ be symmetric matrices. The following conditions are equivalent:*

(i) *$\zeta^T T \zeta < 0$ for all $\zeta \neq 0$ such that $\zeta^T S \zeta \leq 0$.*

(ii) *There exists $\tau > 0$ such that $T - \tau S < 0$.*

When there exists R such that $S = RR^T$ holds in Lemma 2.3, an alternative LMI condition are derived. After defining a notation of matrices, we state this fact, namely, Finsler's lemma [8, 30].

Definition 2.1 For a given matrix $M \in \Re^{n\times m}$, $\tilde{M}$ is a matrix satisfying

$$\tilde{M} \in \Re^{n\times(n-r)}, \quad \tilde{M}^T M = 0, \quad \tilde{M}^T \tilde{M} > 0,$$

where $r := \operatorname{rank} M$.

Lemma 2.4 (Finsler's lemma) *Let $R \in \Re^{n\times m}$ be a real matrix and $T \in \Re^{n\times n}$ be a real symmetric matrix. The following conditions are equivalent:*

(i) *There exists $\mu > 0$ such that $T - \mu RR^T < 0$.*

(ii) $\tilde{R}^T T \tilde{R} < 0$.

2.1.3 Elimination Lemmas

The following lemma is useful for eliminating variables in certain matrix inequalities [8].

Lemma 2.5 *Let G be a real symmetric matrix and U and V be real matrices. Then the following conditions are equivalent:*

(i) *There exists X such that*

$$G + UXV^T + VX^TU^T < 0.$$

(ii)

$$\tilde{U}^T G \tilde{U} < 0, \quad \tilde{V}^T G \tilde{V} < 0. \tag{2.3}$$

We can express (2.3) in another form using Finsler's lemma:

Lemma 2.6 *Let G be a real symmetric matrix and U and V be real matrices. Then the following conditions are equivalent:*

(i) *There exists X such that*

$$G + UXV^T + VX^TU^T < 0.$$

(ii)

$$G - \mu UU^T < 0, \quad G - \mu VV^T < 0$$

holds for some $\mu \in \Re$.

2.2 Properties and a Simple Example of LMIs and SDPs

2.2.1 Properties of LMIs and SDPs

We now discuss some properties of an SDP:

$$\min_x \{c^T x \mid F(x) := F_0 + \sum_{i=1}^m x_i F_i \geq 0\}. \tag{2.4}$$

Here, the problem data are the vector $c \in \Re^m$ and $m+1$ symmetric matrices $F_0, \ldots, F_m \in \Re^{n \times n}$. If $F(x) \geq 0$ and $F(y) \geq 0$, then, for all λ, $0 \leq \lambda \leq 1$,

$$F(\lambda x + (1-\lambda)y) = \lambda F(x) + (1-\lambda)F(y) \geq 0.$$

Thus the constraint of the SDP (2.4) is convex, and therefore, the SDP (2.4) is a convex optimization problem. We see from this fact that every locally optimal solution of an SDP is globally optimal, and we can also regard the optimal solution of an SDP as the solution of the control problem reduced to the SDP. From the viewpoint of numerical computation, an SDP is an important class of convex optimization problems for which readily computable self-concordant barrier functions are known, and therefore, interior-point methods are applicable [42, 39].

Strictly speaking, the constraint of the SDP (2.4)

$$F_0 + \sum_{i=1}^m x_i F_i \geq 0$$

is an affine matrix inequality (AMI). A connection between AMIs and linear matrix inequalities of the form

$$\sum_{i=1}^m x_i F_i \geq 0$$

is discussed in a more general BMI framework in the next section.

2.2.2 A Simple Example of LMIs in Control

Consider the state-feedback stabilization problem for the linear time-invariant system:

$$\dot{x} = Ax + Bu, \tag{2.5}$$

where $x \in \Re^n$ is the system state and $u \in \Re^m$ is the system input. Here A and B are appropriately dimensioned real constant matrices. The closed-loop system with state-feedback $u = Kx$ is as follows:

$$\dot{x} = (A + BK)x. \tag{2.6}$$

System (2.5) is said to be *stabilizable* (via linear state-feedback) if there exists a state-feedback gain K such that the closed-loop system (2.6) is stable, i.e., all the eigenvalues of $A + BK$ are in the open left half plane. From the Lyapunov theorem [8], the closed-loop system (2.6) is stable if and only if there exists $P > 0$ such that

$$(A + BK)^T P + P(A + BK) < 0,$$

or equivalently, there exists $Q > 0$ such that

$$Q(A + BK)^T + (A + BK)Q < 0. \tag{2.7}$$

These conditions are BMIs in K and P or Q, but by a simple change of variables we can obtain an equivalent LMI condition. Define $Y = KQ$, so that for $Q > 0$ we have $K = YQ^{-1}$. Substituting this into (2.7) yields

$$AQ + QA^T + BY + Y^T B^T < 0, \tag{2.8}$$

which is an LMI in Q and Y. Thus system (2.5) is stabilizable if and only if there exists $Q > 0$ and Y such that LMI (2.8) holds. If this LMI is feasible, then function $V(\xi) = \xi^T Q^{-1}\xi$ assures stability of system (2.5) with state-feedback $u = YQ^{-1}x$.

An alternative equivalent condition for stabilizability, involving fewer variables, can be derived using the elimination lemma. From Lemma 2.6, LMI (2.8) holds if and only if there exist $Q > 0$ and a scalar μ such that

$$AQ + QA^T - \mu BB^T < 0. \tag{2.9}$$

Since we can always assume $\mu > 0$ in LMI (2.9), and since this LMI is homogeneous in Q and μ, we can take $\mu = 1$ without loss of generality, thus reducing the number of variables by one. If $Q > 0$ satisfies the LMI (2.9), a stabilizing state-feedback gain is given by

$$K = -\frac{\mu}{2}B^T Q^{-1}.$$

From the elimination lemma (Lemma 2.5), another equivalent condition is

$$\tilde{B}^T(AQ + QA^T)\tilde{B} < 0. \tag{2.10}$$

For any $Q > 0$ satisfying (2.10), a stabilizing state-feedback gain is

$$K = -\frac{\mu}{2}B^T Q^{-1},$$

where μ is any scalar such that (2.9) holds (condition (2.10) implies that such a scalar exists) [8].

2.3 Properties and a Simple Example of BMIs

2.3.1 Properties of BMIs

We will now discuss some properties of a BMI:

$$F(x,y) := F_{00} + \sum_{i=1}^{n_x} x_i F_{i0} + \sum_{j=1}^{n_y} y_j F_{0j} + \sum_{i=1}^{n_x}\sum_{j=1}^{n_y} x_i y_j F_{ij} > 0, \tag{2.11}$$

The variables are $x \in \Re^{n_x}$ and $y \in \Re^{n_y}$. The coefficient matrices $F_{ij} \in \Re^{m\times m}, i = 1,\ldots,n_x, j = 1,\ldots,n_y$ are symmetric.

There are certain cases where the solution to BMI (2.11) is trivial, or where the BMI reduces to an AMI. For example, if $F_{ij} = 0$ for all $i = 1,\ldots,n_x$, $j = 1,\ldots,n_y$, then $F(x,y) > 0$ is in fact an AMI in the pair (x,y). Also, obtaining a feasible solution for the BMI (2.11) is trivial whenever any one of the F_{ij} except F_{00} is positive or negative definite. Further, note that if any of the AMIs in x, i.e., $F_{00} + \sum_{i=1}^{n_x} x_i F_{i0} > 0$, $F_{00} + F_{0j} + \sum_{i=1}^{n_x} x_i(F_{i0} + F_{ij}) > 0$, $j = 1,\ldots,n_y$, has a solution, then the solution to the BMI problem (2.11) trivially follows. The same holds for the corresponding AMIs in y.

If BMI (2.11) is linear in at least one of the variables, say x, then $F(x,y) > 0$ if and only if $F(\kappa x, y) > 0$ for all $\kappa > 0$, i.e., the feasible set of (2.11) will be unbounded if it is non-empty and $\max \lambda_{\min}(F(x,y))$ is either unbounded or negative.

Further, if the BMI is actually linear in both x and y, then $F(\bar{x},\bar{y}) > 0$ if and only if $F(\kappa_x\bar{x}, \kappa_y\bar{y}) > 0$, for all $\kappa_x\kappa_y > 0$. Hence if the feasible set is non-empty, it is unbounded.

We now investigate the underlying geometry of the problem of finding a feasible solution for a BMI. To this end, we define biconvexity of a set and a function, and describe some properties concerning them.

Consider a set $\mathcal{X} \times \mathcal{Y} \subset \Re^{n_x} \times \Re^{n_y}$, where $\mathcal{X}$ is convex in $\Re^{n_x}$ and $\mathcal{Y}$ is convex in $\Re^{n_y}$. Define the x and *y-sections* of $\mathcal{B}$ as follows: $\mathcal{B}_x := \{y \in \mathcal{Y} : (x,y) \in \mathcal{B}\}$ and $\mathcal{B}_y := \{x \in \mathcal{X} : (x,y) \in \mathcal{B}\}$.

Definition 2.2 (Biconvexity of a Set) The set $\mathcal{B} \subset \mathcal{X} \times \mathcal{Y}$ is biconvex if $\mathcal{B}_x$ is convex for every $x \in \mathcal{X}$ and $\mathcal{B}_y$ is convex for every $y \in \mathcal{Y}$.

Proposition 2.1 *The set $\mathcal{B} \subset \mathcal{X} \times \mathcal{Y}$ is biconvex if and only if for every quadruple $(x_1,y_1),(x_1,y_2),(x_2,y_1),(x_2,y_2) \in \mathcal{B}$,*

$$(x_\beta, y_\gamma) := ((1-\beta)x_1 + \beta x_2, (1-\gamma)y_1 + \gamma y_2) \in \mathcal{B}$$

holds for every $(\beta,\gamma) \in [0,1] \times [0,1]$.

A biconvex set is not necessarily convex. For example, consider the shape "L" on the product space $\Re \times \Re$; this is biconvex but not convex and also the level sets for $f(x,y) = xy < 1$, $f(x,y) = xy < -1$. The latter of which is not even connected but still biconvex.

Definition 2.3 (Biconvexity of a Function) A function $f : \mathcal{X} \times \mathcal{Y} \to \Re$ is *biconvex* in (x, y) if it is convex in x for every fixed $y \in \mathcal{Y}$ and convex in y for every fixed $x \in \mathcal{X}$.

Proposition 2.2 *A function $f(x, y)$ is biconvex over $\mathcal{X} \times \mathcal{Y}$ if and only if for any (x_1, y_1), (x_2, y_2), (x_2, y_1), $(x_2, y_2) \in \mathcal{X} \times \mathcal{Y}$,*

$$f(x_\beta, y_\gamma) \leq (1-\beta)(1-\gamma)f(x_1, y_1) + (1-\beta)\gamma f(x_1, y_2) + \beta(1-\gamma)f(x_2, y_1) + \beta\gamma f(x_2, y_2)$$

for every $(\beta, \gamma) \in [0, 1] \times [0, 1]$, where $(x_\beta, y_\gamma) := ((1-\beta)x_1 + \beta x_2, (1-\gamma)y_1 + \gamma y_2)$.

Proposition 2.2 states that a convex function and a biconvex function have a similar property, i.e., just as one dimensional interpolation always overestimates a convex function, two dimensional interpolation always overestimates a biconvex function. A convex set and a biconvex set have a similar property as follows:

Proposition 2.3 *If $f(x, y)$ is biconvex, then its level sets, $\mathcal{L}_c := \{(x, y) \in \mathcal{X} \times \mathcal{Y} : f(x, y) \leq c\}$, $c \in \Re$, are biconvex for all c.*

While convexity and biconvexity have some similar properties, there is great difference between them. One of the reasons for the study of convexity of functions is that every local minimum is always the global minimum for convex functions. However, convexity is not a necessary condition for the local-global property, and in fact strict quasiconvexity is a sufficient condition for every local minimum to be the global minimum [13]. Unfortunately, biconvexity, as we have defined it, does not yield the local-global property.

Here we will return to BMI (2.11). For the function defined by

$$\Lambda(x, y) := -\lambda_{\min}(F(x, y)),$$

$\Lambda(x, y) < 0$ if and only if $F(x, y) > 0$. Thus we see that BMI (2.11) is solvable by reducing the minimization problem of $\Lambda(x, y)$. For $Lambda(x, y)$, the following theorem holds.

Theorem 2.1 *The function $\Lambda(x, y)$ is biconvex over $\Re^{n_x} \times \Re^{n_y}$.*

Proof The proof follows trivially from the well established fact that AMI/LMIs are convex. ■

Therefore, Propositions 2.2, 2.3 hold for $\Lambda(x, y)$. Note that the BMI problem does not have the local-global property because of the biconvexity of $F(x, y) > 0$.

2.3.2 Connection between Biaffine and Bilinear Cases

We now examine the relationships between the bilinear and biaffine matrix inequalities.

For a given bilinear matrix function

$$\bar{F}(x, y) := \sum_{i=1}^{n_x}\sum_{j=1}^{n_y} x_i y_j F_{ij},$$

a biaffine matrix function

$$F(x, y) := F_{00} + \sum_{i=1}^{n_x} x_i F_{i0} + \sum_{j=1}^{n_y} y_j F_{0j} + \sum_{i=1}^{n_x}\sum_{j=1}^{n_y} x_i y_j F_{ij}$$

is easily generated, where F_{00}, F_{i0} and F_{0j} are all zero matrices. Therefore, we see that bilinear matrix inequality problems are a subset of biaffine matrix inequality problems. The following lemma shows that the converse is also true in the sense that every biaffine matrix inequality problem can be represented as a bilinear matrix inequality problem via a suitable argumentation of the F_{ij} matrices.

Lemma 2.7 *Let*

$$\tilde{F}(\tilde{x}, \tilde{y}) := \operatorname{diag}(\tilde{x}_0\tilde{y}_0, \hat{F}(\tilde{x}, \tilde{y})),$$

where

$$\hat{F}(x, y) := \sum_{i=0}^{n_x}\sum_{j=0}^{n_y} x_i y_j F_{ij}.$$

If a pair $(\tilde{x}, \tilde{y})$ satisfies the bilinear matrix inequality $\tilde{F}(\tilde{x}, \tilde{y}) > 0$, then the pair (x, y) given by $x_i = \frac{1}{\tilde{x}_0}\tilde{x}_i$, $i = 1, \ldots, n_x$, $y_j = \frac{1}{\tilde{y}_0}\tilde{y}_j$ $j = 1, \ldots, n_y$ satisfies the biaffine matrix inequality $F(x, y) > 0$.

Proof By construction, $\tilde{x}_0\tilde{y}_0 \operatorname{diag}(1, F(x, y)) = \tilde{F}(\tilde{x}, \tilde{y})$. Dividing both sides of this equality by the positive scalar $\tilde{x}_0\tilde{y}_0$, we obtain the result. ■

In view of Lemma 2.7, it is clear that every biaffine matrix inequality can be reformulated as a bilinear matrix inequality via a trivial augmentation of the F_{ij} matrices. Thus we conclude that the two are mathematically equivalent. Accordingly, we use terms bilinear matrix inequality and biaffine matrix inequality interchangeably and refer to both as simply BMIs.

2.3.3 A Simple Example of a BMI in Control

We here show a simple example of a BMI. Consider the output feedback stabilization problem for the linear time-invariant system:

$$\begin{aligned} \dot{x} &= Ax + Bu \\ y &= Cx, \end{aligned} \tag{2.12}$$

where $x \in \Re^n$ is the system state, $u \in \Re^m$ is the system input and $y \in \Re^p$ is the system output. Here, A, B and C are appropriately dimensioned real constant matrices. The closed-loop system with the output feedback $u = Ky$ is as follows:

$$\dot{x} = (A + BKC)x. \tag{2.13}$$

From the Lyapunov theorem, the closed-loop system (2.13) is stable, i.e., all the eigenvalues of $A + BKC$ are in the open left half plane if and only if there exists $P > 0$ such that

$$(A + BKC)^T P + P(A + BKC) < 0.$$

Thus the system (2.12) is stabilizable via the output feedback if and only if there exists a symmetric matrix P and a matrix K such that the following BMI holds:

$$\begin{bmatrix} -(A + BKC)^T P - P(A + BKC) & 0 \\ 0 & P \end{bmatrix} > 0. \tag{2.14}$$

Note that this BMI condition (2.14) cannot be converted to an equivalent LMI condition by the techniques described in the previous section.

Chapter 3

Sensitivity Analysis in SDP and Its Application to Control Systems with Uncertainty

A general objective in control system design is to make the designed closed-loop system stable and achieve a specified control performance. However, in the case where a control system includes uncertain parameters with estimated errors and/or changes, stability and performance are subject to these parameters. It is, therefore, important to obtain sensitivity information of stability and control performance with respect to such uncertain parameters. The sensitivity information is also useful for determining physical parameters such as damping and stiffness constants from the viewpoint of structural design [43].

The goal of this chapter is to give an analysis method for estimating the sensitivity of control performance via SDP, based on the results in [58, 59]. We first provide some results concerning sensitivity analysis of SDP and then apply them to control systems.

3.1 Sensitivity Analysis of SDP

As stated in the previous chapter, SDP is a special form of convex programming, and so far, various properties of SDP have been studied from the theoretical and practical viewpoints [56, 34]. The optimality of SDP is one of the most important properties and has been exploited to develop efficient interior-point algorithms. However, sensitivity analysis of SDP has been less studied until now. For convex (nonlinear) programming, it is well known that optimality and the so-called implicit function theorem play important roles in sensitivity analysis [11]. However the results of sensitivity analysis in nonlinear programming in [11] cannot be readily applied to SDP since the constraint of an SDP is in a matrix form but not a vector form.

To cope with the above difficulty, we first propose a new form of complementarity condition. Using the proposed complementarity condition, we provide results concerning sensitivity analysis, i.e, a technique for estimating the sensitivity of the optimal solution and the optimal value function with respect to perturbations in SDP.

3.1.1 Optimality of SDP

We first present well-known results of optimality of SDP in order to prepare for sensitivity analysis of SDP. We then propose a new form of complementarity condition such that the Jacobian matrix of a vector equality related to optimality conditions is a square matrix. As a result, we can exploit the implicit function theorem in sensitivity analysis.

Consider the SDP (denoted (P_θ) hereafter) with an uncertain parameter:

$$\psi_p(\theta) := \min_x c^T x$$

subject to

$$F(x,\theta) := F_0(\theta) + \sum_{j=1}^{m} x_j F_j(\theta) \geq 0,$$

where $x \in \Re^m$ is a vector variable, $c \in \Re^m$ is a constant vector and $F_j(\theta) \in \Re^{n\times n}, j = 0,\ldots,m$ are symmetric matrix functions of an uncertain parameter $\theta \in \Re^l$. We assume that $F_j(\theta)$ is continuously differentiable in θ. As seen in the previous chapter, the SDP (P_θ) is convex for any fixed θ. We are interested in analyzing the behavior of an optimal value function of the SDP (P_{θ_0}) when a fixed parameter θ_0 is subject to perturbation. First, we deal with the SDP (P_{θ_0}) with the fixed parameter θ_0. The dual problem (denoted (D_{θ_0}) hereafter) associated with the SDP (P_{θ_0}) is

$$\psi_d(\theta_0) := \max_y(-\,\mathbf{Tr}\, F_0(\theta_0)G(y))$$

subject to

$$\mathbf{Tr}\, F_j(\theta_0)G(y) = c_j, \quad j = 1,\ldots,m$$

and

$$G(y) = \sum_{j=1}^{q} y_j G_j \geq 0,$$

where $q := n(n+1)/2$, $y \in \Re^q$ is a vector variable and $G_j \in \Re^{n\times n}, j = 1,\ldots,q$ are symmetric and linearly independent, i.e., the system of homogeneous linear equations $\sum_{j=1}^{q} y_j G_j = 0$ in q variables $y_1, y_2, \ldots, y_q$ admits no nontrivial solution, and $\mathbf{Tr}\, A$ denotes the trace of a matrix A.

Throughout this chapter, we make the following assumption for the SDPs (P_{θ_0}) and (D_{θ_0}).

Assumption 3.1

1. $F_j(\theta_0)$, $j = 1,\ldots,m$ are linearly independent.

2. There exists an interior and feasible solution (x_0, y_0) of the SDPs (P_{θ_0}) and (D_{θ_0}) as follows:
$$F(x_0,\theta_0) > 0,\ G(y_0) > 0,$$
$$\mathbf{Tr}\, F_j(\theta_0)G(y_0) = c_j, \quad j = 1,\ldots,m.$$

The assumption 1 is necessary for solving SDP via a type of primal-dual interior-point method. The assumption 2 guarantees that both SDPs (P_{θ_0}) and (D_{θ_0}) have optimal solutions [53, 34].

For the SDPs (P_{θ_0}) and (D_{θ_0}), optimality conditions can be stated as follows [56, 34].

Theorem 3.1 *Under Assumption 3.1, (x^*, y^*) is an optimal solution of the SDPs (P_{θ_0}) and (D_{θ_0}) if and only if*

$$F(x^*,\theta_0) = F_0(\theta_0) + \sum_{j=1}^{m} x_j^* F_j(\theta_0) \geq 0, \tag{3.1}$$

$$G(y^*) = \sum_{j=1}^{q} y_j^* G_j \geq 0, \tag{3.2}$$

$$\mathbf{Tr}\, F_j(\theta_0)G(y^*) = c_j, \quad j = 1,\ldots,m, \tag{3.3}$$

$$F(x^*,\theta_0)G(y^*) + G(y^*)F(x^*,\theta_0) = 0. \tag{3.4}$$

Then

$$\psi_p(\theta_0) = \psi_d(\theta_0).$$

Namely, the primal optimal value coincides with the dual optimal value.

Remark 3.1 In sensitivity analysis of nonlinear programming [11], the implicit function theorem is applied to a function related to equality conditions in optimality conditions. In order to take the same approach, the Jacobian matrix of the function with respect to uncertain parameters should be square. To this end, we propose the new type of complementarity condition (3.3).

The forms of $\mathbf{Tr}\, F(x^*, \theta_0)G(y^*) = 0$ or $F(x^*, \theta_0)G(y^*) = 0$ are usually used as complementarity conditions. If the primal feasibility (3.1) and the dual feasibility (3.2), (3.3) are both satisfied, the following lemma shows that these complementarity conditions are equivalent.

Lemma 3.1 *For any symmetric positive semidefinite matrices F, $G \in \Re^{n\times n}$, the following conditions are equivalent:*

(i) $\mathbf{Tr}\, FG = 0$.

(ii) $FG = 0$.

(iii) $FG + GF = 0$.

Proof It is easy to see that (ii) $\Longrightarrow$ (iii) and (iii) $\Longrightarrow$ (i) hold. Hence we have only to show (i) $\Longrightarrow$ (ii) to prove the equivalence among (i), (ii) and (iii).

Since F and G are positive semidefinite, there exist positive semidefinite matrices $F^{\frac{1}{2}}$, $G^{\frac{1}{2}}$ such that

$$F = F^{\frac{1}{2}}F^{\frac{1}{2}}, \quad G = G^{\frac{1}{2}}G^{\frac{1}{2}}.$$

Thus, when condition (i), i.e., $\mathbf{Tr}\, FG = 0$ holds, we have

$$\mathbf{Tr}\, F^{\frac{1}{2}}G^{\frac{1}{2}}G^{\frac{1}{2}}F^{\frac{1}{2}} = 0.$$

Since $F^{\frac{1}{2}}G^{\frac{1}{2}}G^{\frac{1}{2}}F^{\frac{1}{2}}$ is positive semidefinite, we have

$$F^{\frac{1}{2}}G^{\frac{1}{2}}G^{\frac{1}{2}}F^{\frac{1}{2}} = 0.$$

Furthermore, since

$$F^{\frac{1}{2}}G^{\frac{1}{2}}G^{\frac{1}{2}}F^{\frac{1}{2}} = (F^{\frac{1}{2}}G^{\frac{1}{2}})(F^{\frac{1}{2}}G^{\frac{1}{2}})^T,$$

we obtain $F^{\frac{1}{2}}G^{\frac{1}{2}} = 0$, and therefore

$$FG = F^{\frac{1}{2}}F^{\frac{1}{2}}G^{\frac{1}{2}}G^{\frac{1}{2}} = 0.$$

This completes the proof. ∎

The complementarity conditions of SDP give an eigenstructure between F and G as shown in the following lemma. This fact plays an important role in proving the results of sensitivity analysis in the next subsection.

Lemma 3.2 *Suppose that the conditions in Lemma 3.1 are satisfied. Then there exists an orthogonal matrix U such that*

$$\begin{aligned} U^TFU &= \begin{bmatrix} \Lambda_F & 0 \\ 0 & 0 \end{bmatrix}, \quad \Lambda_F = \mathrm{diag}(\lambda_1(F), \ldots, \lambda_{r_F}(F)), \\ U^TGU &= \begin{bmatrix} 0 & 0 \\ 0 & \Lambda_G \end{bmatrix}, \quad \Lambda_G = \mathrm{diag}(\lambda_1(G), \ldots, \lambda_{r_G}(G)), \end{aligned}$$

where $r_F := \mathrm{rank}(F)$, $r_G := \mathrm{rank}(G)$, and $\lambda_i(F)$, $i = 1, \ldots, r_F$ and $\lambda_j(G)$, $j = 1, \ldots, r_G$ are all positive eigenvalues of F and G, respectively. Also, $r_G \le n - r_F$ holds.

Proof Since F is positive semidefinite, there exists an orthogonal matrix U_F such that

$$U_F^TFU_F = \begin{bmatrix} \Lambda_F & 0 \\ 0 & 0 \end{bmatrix}. \tag{3.5}$$

From condition (ii) in Lemma 3.1, i.e., $FG = GF = 0$,

$$(U_F^TFU_F)(U_F^TGU_F) = (U_F^TGU_F)(U_F^TFU_F) = 0. \tag{3.6}$$

Then define $U_F^TGU_F =: [g_{ij}]$, $i, j = 1, \ldots, n$ and we can rewrite (3.6) as

$$\left[\begin{array}{ccc} \lambda_1 g_{11} & \cdots & \lambda_1 g_{1n} \\ \vdots & \cdots & \vdots \\ \lambda_{r_F} g_{r_F 1} & \cdots & \lambda_{r_F} g_{r_F n} \\ \hline & O_{(n-r_F)\times n} & \end{array}\right] = \left[\begin{array}{ccc|c} \lambda_1 g_{11} & \cdots & \lambda_{r_F} g_{1r_F} & \\ \vdots & \cdots & \vdots & O_{n\times(n-r_F)} \\ \lambda_1 g_{n1} & \cdots & \lambda_{r_F} g_{nr_F} & \end{array}\right] = 0.$$

From $\lambda_i > 0$, $i = 1, \ldots, r_F$, it follows that $g_{ij} = 0$, $i \le r_F$ or $j \le r_F$, and we obtain

$$U_F^T G U_F = \begin{bmatrix} 0 & 0 \\ 0 & \tilde{G} \end{bmatrix},$$

where $\tilde{G} \in \Re^{(n-r_F)\times(n-r_F)}$ is positive semidefinite. Hence $U_F^T G U_F$ can be diagonalized by a block-diagonal orthogonal matrix

$$U_G := \begin{bmatrix} I_{r_F} & 0 \\ 0 & U_{\tilde{G}} \end{bmatrix},$$

so that

$$U_G^T (U_F^T G U_F) U_G = \begin{bmatrix} 0 & 0 \\ 0 & \Lambda_G \end{bmatrix}.$$

where $r_G \le n - r_F$. Then using an orthogonal matrix defined by $U := U_F U_G$, we see that F and G are simultaneously diagonalized by U as follows:

$$U^T F U = \begin{bmatrix} \Lambda_F & 0 \\ 0 & 0 \end{bmatrix},$$

$$U^T G U = \begin{bmatrix} 0 & 0 \\ 0 & \Lambda_G \end{bmatrix}.$$

■

Remark 3.2 If the optimality conditions (3.1)–(3.4) are satisfied for the SDPs (P_{θ_0}) and (D_{θ_0}), we say that the *strict complementarity condition* holds if $\mathrm{rank}(F(x^*, \theta_0)) + \mathrm{rank}(G(y^*)) = n$ holds.

3.1.2 Sensitivity Analysis of SDP

In this subsection, we give fundamental accounts on sensitivity analysis of SDP based on the results in [58, 59]. This results are applied to control systems with uncertainty in the next section, so that we can evaluate sensitivity of control performance.

Before we begin to state the results of sensitivity analysis of SDP, we define the following function Φ associated with (3.3) and (3.4):

$$\Phi(x, y, \theta) = \begin{bmatrix} \mathbf{Tr}\, F_1(\theta)G(y) - c_1 \\ \vdots \\ \mathbf{Tr}\, F_m(\theta)G(y) - c_m \\ \mathbf{Vec}(F(x,\theta)G(y) + G(y)F(x,\theta)) \end{bmatrix}, \tag{3.7}$$

where for a symmetric matrix $A \in \Re^{n\times n}$:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix},$$

$\mathbf{Vec}(A) \in \Re^{n(n+1)/2}$ is defined as follows:

$$\mathbf{Vec}(A) := [a_{11}, a_{21}, \ldots, a_{n1}, a_{22}, \ldots, a_{n2}, a_{33}, \ldots, a_{nn}]^T,$$

i.e., **Vec** stacks the rows of A from the principal diagonal downwards in a column vector. We are now ready to evaluate the variation of the optimal value functions $\psi_p(\theta)$ and $\psi_d(\theta)$ with respect to a variation in the parameter θ. From now on, $\nabla_\theta x$ denotes the $m\times l$ Jacobi matrix whose (i,j) element is $\partial x_i(\theta)/\partial\theta_j$. The following theorem is our main result.

Theorem 3.2 *For SDPs* (P_{θ_0}) *and* (D_{θ_0}), *suppose that the following conditions are satisfied:*

(i) (x^*, y^*) *is an optimal solution of the SDPs* (P_{θ_0}) *and* (D_{θ_0}) *(i.e., (3.1)–(3.4) in Theorem 3.1 hold);*

(ii) $\mathrm{rank}(F(x^*, \theta_0)) + \mathrm{rank}(G(y^*)) = n$ *holds, (i.e., the strict complementarity condition holds);*

(iii) *The* $(m+q)\times(m+q)$ *matrix*

$$[\nabla_x \Phi(x^*, y^*, \theta_0) \quad \nabla_y \Phi(x^*, y^*, \theta_0)]$$

is nonsingular (i.e., the Jacobian matrix of Φ *with respect to* (x,y) *is nonsingular).*

Then, there exist a neighborhood Ω of θ_0 and continuously differentiable functions $x(\theta): \Omega \to \Re^m$ and $y(\theta): \Omega \to \Re^q$ satisfying the following conditions for all $\theta \in \Omega$:

(a) $x(\theta_0) = x^*$, $y(\theta_0) = y^*$;

(b) $x(\theta)$ *and* $y(\theta)$ *satisfy the above assumptions (i)–(iii) for the SDPs* (P_θ) *and* (D_θ);

(c)

$$\begin{bmatrix} \nabla_\theta x(\theta) \\ \nabla_\theta y(\theta) \end{bmatrix} = - \left[\nabla_x \Phi(x(\theta), y(\theta), \theta) \quad \nabla_y \Phi(x(\theta), y(\theta), \theta)\right]^{-1} \nabla_\theta \Phi(x(\theta), y(\theta), \theta); \tag{3.8}$$

(d)

$$\begin{aligned} \nabla_\theta \psi_p(\theta) &= \nabla_\theta \psi_d(\theta), \\ \nabla_\theta \psi_p(\theta) &= c^T \nabla_\theta x(\theta), \end{aligned} \tag{3.9}$$

$$\nabla_\theta \psi_d(\theta) = \begin{bmatrix} -\mathbf{Tr}\left(\frac{\partial F_0}{\partial \theta_1}(\theta) G(y(\theta)) + F_0(\theta) G(\frac{\partial y}{\partial \theta_1}(\theta))\right) \\ \vdots \\ -\mathbf{Tr}\left(\frac{\partial F_0}{\partial \theta_l}(\theta) G(y(\theta)) + F_0(\theta) G(\frac{\partial y}{\partial \theta_l}(\theta))\right) \end{bmatrix}^T. \tag{3.10}$$

Proof

(a) Under the assumption (iii), using the implicit function theorem [11] with respect to (3.7), we can conclude that there exist a neighborhood Ω of θ_0 and continuously differentiable functions $x(\theta): \Omega \to \Re^m$ and $y(\theta): \Omega \to \Re^q$ satisfying (a) and

$$\Phi(x(\theta), y(\theta), \theta) = 0 \quad \text{for all } \theta \in \Omega. \tag{3.11}$$

(b) According to Lemma 3.2 and the assumption (i), (ii), there exists an orthogonal matrix U_0 such that

$$\begin{aligned} U_0^T F(x(\theta_0), \theta_0) U_0 &= \begin{bmatrix} \Lambda_{F0} & 0 \\ 0 & 0 \end{bmatrix}, \\ U_0^T G(y(\theta_0)) U_0 &= \begin{bmatrix} 0 & 0 \\ 0 & \Lambda_{G0} \end{bmatrix}, \end{aligned}$$

where $\Lambda_{F0} \in \Re^{r_F \times r_F}$ and $\Lambda_{G0} \in \Re^{r_G \times r_G}$ are diagonal and positive definite, and $r_F + r_G = n$ holds. Since $F(x(\theta), \theta)$ is symmetric for every θ, and the eigenvalues of $F(x(\theta), \theta)$ and $G(y(\theta))$ are continuous functions of θ, there exist an orthogonal matrix $U(\theta)$ and a sufficiently small neighborhood Ω such that

$$U_F(\theta)^T F(x(\theta), \theta) U_F(\theta) = \begin{bmatrix} \Lambda_F(\theta) & 0 \\ 0 & A(\theta) \end{bmatrix}, \tag{3.12}$$

$$U_F(\theta)^T G(y(\theta)) U_F(\theta) = \begin{bmatrix} B(\theta) & C(\theta) \\ C(\theta)^T & D(\theta) \end{bmatrix}, \tag{3.13}$$

$$\Lambda_F(\theta) > 0, \quad D(\theta) > 0, \quad \text{for all } \theta \in \Omega, \tag{3.14}$$

$$U_F(\theta_0) = U_0, \quad \Lambda_F(\theta_0) = \Lambda_{F0}, \quad D(\theta_0) = \Lambda_{G0},$$

where $\Lambda_F(\theta) = \mathrm{diag}(\lambda_1, \ldots, \lambda_{r_F})$, $A(\theta) = \mathrm{diag}(a_1, \ldots, a_{r_G})$, $B(\theta) = [b_{ij}] \in \Re^{r_F \times r_F}$, $C(\theta) \in \Re^{r_F \times r_G}$ and $D(\theta) = [d_{ij}] \in \Re^{r_G \times r_G}$. Note that $U_F(\theta)$ diagonalize $F(x(\theta), \theta)$. The satisfaction of (3.11) means that, for θ near θ_0,

$$\mathbf{Tr}\, F_j(\theta) G(y(\theta)) = c_j, \quad j = 1, \ldots, m \tag{3.15}$$

$$F(x(\theta), \theta) G(y(\theta)) + G(y(\theta)) F(x(\theta), \theta) = 0. \tag{3.16}$$

From (3.12), (3.13) and (3.16), we have

$$\begin{bmatrix} \Lambda_F(\theta) & 0 \\ 0 & A(\theta) \end{bmatrix} \begin{bmatrix} B(\theta) & C(\theta) \\ C(\theta)^T & D(\theta) \end{bmatrix} + \begin{bmatrix} B(\theta) & C(\theta) \\ C(\theta)^T & D(\theta) \end{bmatrix} \begin{bmatrix} \Lambda_F(\theta) & 0 \\ 0 & A(\theta) \end{bmatrix} = 0, \tag{3.17}$$

that is,

$$\begin{aligned} \Lambda_F(\theta) B(\theta) + B(\theta) \Lambda_F(\theta) &= 0, & (3.18) \\ \Lambda_F(\theta) C(\theta) + C(\theta) A(\theta) &= 0, & (3.19) \\ A(\theta) D(\theta) + D(\theta) A(\theta) &= 0. & (3.20) \end{aligned}$$

We will now observe (3.18)–(3.20) to show that $A(\theta) = 0$, $B(\theta) = 0$ and $C(\theta) = 0$ for $\theta \in \Omega$. From the (i, j) element of each side in (3.18),

$$\lambda_i b_{ij} + b_{ij} \lambda_j = 0. \tag{3.21}$$

Since $\lambda_i + \lambda_j > 0$, it follows that $b_{ij} = 0$, i.e., $B(\theta) = 0$. From the (i, j) element in (3.20),

$$a_i d_{ii} + d_{ii} a_i = 0. \tag{3.22}$$

Since $D(\theta) > 0$, we see that $d_{ii} > 0$. Thus we get $a_i = 0$, i.e., $A(\theta) = 0$. From (3.19), notice that $A(\theta) = 0$ and $\Lambda_F(\theta) > 0$, so that $C(\theta) = 0$. Consequently, we obtain for $\theta \in \Omega$

$$U_F(\theta)^T F(x(\theta), \theta) U_F(\theta) = \begin{bmatrix} \Lambda_F(\theta) & 0 \\ 0 & 0 \end{bmatrix}, \tag{3.23}$$

$$U_F(\theta)^T G(y(\theta)) U_F(\theta) = \begin{bmatrix} 0 & 0 \\ 0 & \Lambda_G(\theta) \end{bmatrix}. \tag{3.24}$$

Furthermore, in the same way as in the proof of Lemma 3.2, we can show that there exists an orthogonal matrix $T(\theta)$ which diagonalizes $F(x(\theta), \theta)$ and $G(x(\theta), \theta)$ simultaneously as follows:

$$U(\theta)^T F(x(\theta), \theta) U(\theta) = \begin{bmatrix} \Lambda_F(\theta) & 0 \\ 0 & 0 \end{bmatrix}, \quad \Lambda_F(\theta) > 0,$$

$$U(\theta)^T G(y(\theta)) U(\theta) = \begin{bmatrix} 0 & 0 \\ 0 & \Lambda(\theta) \end{bmatrix}, \quad \Lambda_G(\theta) > 0,$$

where $\Lambda_G(\theta) \in \Re^{r_G \times r_G}$. Hence, for $\theta \in \Omega$, the following hold (i.e., assumptions (i) and (ii) hold for $\theta \in \Omega$):

$$F(x(\theta), \theta) = F_0(\theta) + \sum_{j=1}^{m} x_j(\theta) F_j(\theta) \geq 0,$$

$$G(y(\theta)) = \sum_{j=1}^{q} y_j(\theta) G_j \geq 0,$$

$$\mathrm{rank}(F(x(\theta), \theta)) + \mathrm{rank}(G(y(\theta))) = n.$$

Moreover, $[\nabla_x \Phi \quad \nabla_y \Phi]$ is nonsingular for $\theta \in \Omega$ (i.e., the assumption (iii) holds for $\theta \in \Omega$) when Ω is sufficiently small, since $\Phi(x(\theta), y(\theta), \theta)$ is continuously differentiable in θ

(c) $\Phi(x(\theta), y(\theta), \theta) = 0$ can be differentiated with respect to θ to yield explicit expressions of the first partial derivatives of the vector function $\Phi(x(\theta), y(\theta), \theta)$. It follows that

$$[\nabla_x \Phi \quad \nabla_y \Phi] \begin{bmatrix} \nabla_\theta x(\theta) \\ \nabla_\theta y(\theta) \end{bmatrix} + \nabla_\theta \Phi = 0. \tag{3.25}$$

Since $[\nabla_x \Phi \quad \nabla_y \Phi]$ is nonsingular for $\theta \in \Omega$, we obtain (3.8).

(d) Since $x(\theta)$ and $y(\theta)$ are continuously differentiable in θ, we see that the optimal value functions $\psi_p(\theta)$ and $\psi_d(\theta)$ are continuously differentiable in θ. Thus (3.9) and (3.10) hold for $\theta \in \Omega$.

■

We show how to calculate the sensitivity in the case where $\theta \in \Omega$ is fixed to be θ_0. In this case, (c) in Theorem 3.2 is as follows:

$$\begin{bmatrix} \nabla_\theta x(\theta_0) \\ \nabla_\theta y(\theta_0) \end{bmatrix} = -\left[\nabla_x \Phi(x^*, y^*, \theta_0) \; \nabla_y \Phi(x^*, y^*, \theta_0)\right]^{-1} \nabla_\theta \Phi(x^*, y^*, \theta_0),$$

where (x^*, y^*) is the optimal solution to the SDPs (P_{θ_0}) and (D_{θ_0}). For notational convenience, we define $F^* := F(x^*, \theta_0)$ and $G^* := G(y^*)$. Then $\nabla_x \Phi(x^*, y^*, \theta_0)$, $\nabla_y \Phi(x^*, y^*, \theta_0)$ and $\nabla_\theta \Phi(x^*, y^*, \theta_0)$ are calculated as follows:

$$\nabla_x \Phi(x^*, y^*, \theta_0) = \left[\begin{array}{ccc} \multicolumn{3}{c}{O_{m \times m}} \\ \hline \mathbf{Vec}\,(F_1(\theta_0) G^* + G^* F_1(\theta_0)) & \cdots & \mathbf{Vec}\,(F_m(\theta_0) G^* + G^* F_m(\theta_0)) \end{array}\right], \tag{3.26}$$

$$\nabla_y \Phi(x^*, y^*, \theta_0) = \left[\begin{array}{ccc} \mathbf{Tr}\, F_1(\theta_0) G_1 & \cdots & \mathbf{Tr}\, F_1(\theta_0) G_q \\ \vdots & \cdots & \vdots \\ \mathbf{Tr}\, F_m(\theta_0) G_1 & \cdots & \mathbf{Tr}\, F_m(\theta_0) G_q \\ \hline \mathbf{Vec}\,(F^* G_1 + G_1 F^*) & \cdots & \mathbf{Vec}\,(F^* G_q + G_q F^*) \end{array}\right], \tag{3.27}$$

$$\nabla_\theta \Phi(x^*, y^*, \theta_0) = \left[\begin{array}{ccc} \mathbf{Tr}\, \frac{\partial F_1(\theta_0)}{\partial \theta_1} G^* & \cdots & \mathbf{Tr}\, \frac{\partial F_1(\theta_0)}{\partial \theta_l} G^* \\ \vdots & \cdots & \vdots \\ \mathbf{Tr}\, \frac{\partial F_m(\theta_0)}{\partial \theta_1} G^* & \cdots & \mathbf{Tr}\, \frac{\partial F_m(\theta_0)}{\partial \theta_l} G^* \\ \hline \mathbf{Vec}\left(\frac{\partial F(x^*, \theta_0)}{\partial \theta_1} + G^* \frac{\partial F(x^*, \theta_0)}{\partial \theta_1}\right) & \cdots & \mathbf{Vec}\left(\frac{\partial F(x^*, \theta_0)}{\partial \theta_l} + G^* \frac{\partial F(x^*, \theta_0)}{\partial \theta_l}\right) \end{array}\right]. \tag{3.28}$$

3.2 Application to Control Systems with Uncertainty

It is often desirable to estimate the sensitivity of the control performance against uncertain parameters of a control system. In this section, we describe how to apply the results in the previous section to such control systems. We here deal with the sensitivity of H^∞ norm of the following linear system with an uncertain parameter $\theta \in \Re^l$.

$$\begin{aligned} \dot{x} &= A(\theta)x + B(\theta)w \\ y &= C(\theta)x \end{aligned}$$

where $A(\theta)$, $B(\theta)$, $C(\theta)$ are continuously differentiable in θ. The following lemma holds for the H^∞ norm of the transfer function G_{yw} from w to y [8].

Lemma 3.3 *Suppose A is a stable matrix. Then*

$$\|G_{yw}\|_\infty := \|C(sI - A)^{-1}B\|_\infty < \gamma$$

if and only if there exists $P > 0$ such that

$$\begin{bmatrix} -A^TP - PA - C^TC & -PB \\ -B^TP & \gamma^2 I \end{bmatrix} \geq 0.$$

From Lemma 3.3, we see that $\|G_{yw}(\theta)\|_\infty^2$ is equal to the optimal value of the SDP (P_θ) where the variables are P and γ^2, the objective is γ^2, and the constraint is

$$\begin{bmatrix} -A(\theta)^TP - PA(\theta) - C(\theta)^TC(\theta) & -PB(\theta) & 0 \\ -B(\theta)^TP & \gamma^2 I & 0 \\ 0 & 0 & P \end{bmatrix} \geq 0. \qquad (3.29)$$

Therefore, we can compute the sensitivity of $\|G_{yw}(\theta)\|_\infty^2$ with respect to the parameter θ by using the results in the previous section for the SDP (P_θ)

Since LMI (3.29) has a block-diagonal structure:

$$\begin{bmatrix} F_a(x,\theta) & 0 \\ 0 & F_b(x,\theta) \end{bmatrix} \geq 0, \qquad (3.30)$$

the variable G of the corresponding dual problem should have the same block-diagonal structure as LMI (3.30), i.e., $G = \mathrm{diag}(G_a, G_b)$, from the viewpoint of computational efficiency [55, 56]. Accordingly, Φ in (3.7) is as follows:

$$\Phi(x,y,\theta) = \begin{bmatrix} \mathbf{Tr}\, F_1(\theta)G(y) - c_1 \\ \vdots \\ \mathbf{Tr}\, F_m(\theta)G(y) - c_m \\ \mathbf{Vec}(F_a(x,\theta)G_a(y) + G_a(y)F_a(x,\theta)) \\ \mathbf{Vec}(F_b(x,\theta)G_b(y) + G_b(y)F_b(x,\theta)) \end{bmatrix}.$$

Note that computational efforts for the sensitivity as well as SDP are reduced because the dimension of Φ is reduced.

3.3 Numerical Example

In this section we apply the previous results to a control system with an uncertain parameter and compute the sensitivity of the H^∞ norm of the control system. Consider the linear time-invariant system [44]

$$\begin{aligned} \dot{x} &= Ax + Bw \\ y &= Cx, \end{aligned}$$

where

$$A := \begin{bmatrix} -\frac{D_m}{J_m} & 0 & -\frac{K_s}{J_m} \\ 0 & -\frac{D_l}{J_l} & \frac{K_s}{J_l} \\ 1 & -1 & 0 \end{bmatrix}, \quad B := \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix},$$

$$C := \begin{bmatrix} 0 & 1 & 0 \end{bmatrix},$$

and the parameter $\theta := [J_l, J_m, D_l, D_m, K_s]^T$ takes the nominal value $\theta_0 = [1, 1, 1, 0.1, 5]^T$. Then, using the primal-dual interior-point method [56] to compute the H^∞ norm of the transfer function $G_{yw}(\theta_0)$ from w to y, we obtain

$$\|G_{yw}(\theta_0)\|_\infty = 0.9109.$$

Table 3.1: Sensitivity of H^∞ norm

	J_l	J_m	D_l	D_m	K_s
Proposed method	7.4949×10^{-1}	1.5967×10^{-1}	-8.2812×10^{-1}	-7.9334×10^{-1}	-3.4219×10^{-4}
Finite difference method	7.4947×10^{-1}	1.5967×10^{-1}	-8.2809×10^{-1}	-7.9331×10^{-1}	-3.4226×10^{-4}

From the primal and dual optimal solution (x^*, y^*), we can see that conditions (ii) and (iii) in Theorem 3.2 are satisfied since the following conditions hold:

$$\begin{aligned}
&\mathrm{rank}(F(x^*, \theta_0)) = 5,\\
&\mathrm{rank}(G(y^*)) = 2,\\
&\mathrm{rank}([\nabla_x \Phi(x^*, y^*, \theta_0)\ \nabla_y \Phi(x^*, y^*, \theta_0)]) = 23.
\end{aligned}$$

Now, we calculate the sensitivity $\nabla_\theta \|G_{yw}\|_\infty$ in two ways. First, we use the method of finite differences. Using the SDP software of [55], we calculate H^∞ to an accuracy of 1.0×10^{-10}. We assume a perturbation size of 1.0×10^{-5}. Next, we calculate the sensitivity by using the proposed method.

Table 3.1 shows the results. Note that both sensitivity results calculated by the finite difference method and the proposed method are almost the same, but the proposed method gives the exact sensitivity value. Also note that the H^∞ norm of the above control system is insensitive with respect to K_s.

Chapter 4

Control System Design Considering a Tradeoff between Uncertainty and Performance

This chapter presents a design method of control systems in such a way that a designer can flexibly take account of tradeoffs between evaluated uncertainty ranges and the level of control performance. The results of this chapter are based on [63].

Robust control theory has been highly successful in the past two decades. The majority of contributions are however made upon the assumption that we can take a fixed set of uncertainty and then we attempt to give a guaranteed level of performance for all plants that belong to such a prespecified class of uncertainties. In other words, we aim at a worst-case design.

Whilst this leads in general to a very safe design, a possible drawback is that such a class of uncertainty gives a hard bound, and the worst-case design tends to give a rather conservative result. Furthermore, the uncertainty bounds usually employed are deterministic, and no further detailed information is imposed. On the other hand, the uncertainty encountered in practice may be subject to a certain (probabilistic) distribution, and it is possible that we may impose less stringent performance on the plant data that may appear less frequently. The standard robust control theory is not adequate for dealing with such a situation.

In view of this observation, several approaches have been proposed. For example, studies on the average performance [5, 6, 14, 31, 57], and a robust control method with distribution condition [52]. In particular, various randomized algorithms [31, 57] have been proposed as a new, attractive method for dealing with this problem.

We here study however, from a slightly different viewpoint, the tradeoff between the evaluated uncertainty range and control performance, and reduce the design problem to a semi-infinite programming problem with BMI constraints.

4.1 Conservativeness of Standard Robust Performance Control

Let us begin by considering the following example [19]:

$$\frac{d}{dt}x(t) = A(t)x(t) + Bu(t) \tag{4.1}$$

$$A(t) = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -k(t) & k(t) & -f(t) & f(t) \\ 10k(t) & -10k(t) & 10f(t) & -10f(t) \end{bmatrix},$$

$$B = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad x_0 = x(0) = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

$$0.08 \le k(t) \le 0.32, \quad 0.008 \le f(t) \le 0.032. \tag{4.2}$$

Introduce a polytope $\Omega(\theta)$ parametrized by $\theta \in \Re$ as follows:

$$\begin{aligned} \Omega(\theta) \; &= \; \mathbf{Co}\{A_1(\theta), A_2(\theta), A_3(\theta), A_4(\theta)\}, \\ & A_1(\theta) = A_0 + \theta\hat{A}_1, \qquad A_2(\theta) = A_0 + \theta\hat{A}_2, \\ & A_3(\theta) = A_0 + \theta\hat{A}_3, \qquad A_4(\theta) = A_0 + \theta\hat{A}_4, \end{aligned}$$

$$A_0 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -0.2 & 0.2 & -0.02 & 0.02 \\ 2 & -2 & 0.2 & -0.2 \end{bmatrix},$$

$$\hat{A}_1 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -0.12 & 0.12 & -0.012 & 0.012 \\ 1.2 & -1.2 & 0.12 & -0.12 \end{bmatrix}, \qquad \hat{A}_2 = -\hat{A}_1,$$

$$\hat{A}_3 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -0.12 & 0.12 & 0.012 & -0.012 \\ 1.2 & -1.2 & -0.12 & 0.12 \end{bmatrix}, \qquad \hat{A}_4 = -\hat{A}_3.$$

For each fixed θ, $\Omega(\theta)$ gives a set of uncertainty, and we see that

$$A(t) \in \Omega(1) \tag{4.3}$$

holds for all t. Thus the system with parameter variations (4.1), (4.2) can be rewritten as the system with polytopic uncertainty (4.1), (4.3).

Consider the following performance index for system (4.1), (4.3):

$$J(\theta) = \sup_{A(\cdot)\in\Omega(\theta)} \int_0^\infty (x(t)^T Q x(t) + u(t)^T R u(t))dt,$$

where $Q > 0$ and $R > 0$. A straightforward objective is to minimize this functional for each uncertainty set $\Omega(\theta)$. Unfortunately this objective is not easily tractable (partly because of supremum taken over the uncertainty set), so we relax this problem to a minimization problem of an upper bound of the performance index.

Derive an upper bound of the above performance index according to [1] as follows: Suppose that a quadratic functional $\psi^T P \psi$ with $P > 0$ satisfies

$$\frac{d}{dt}x(t)^T P x(t) < -x(t)^T (Q + K^T R K) x(t) \tag{4.4}$$

for all $t \ge 0$, and for all x and u satisfying (4.1) with $x(T) = 0$. Integrating both sides from 0 to T, we obtain

$$x_0^T P x_0 > \int_0^T x(t)^T (Q + K^T R K) x(t) dt,$$

and hence this functional gives a upper bound for $J(\theta)$. Replace the objective with that of the minimization of this functional with state-feedback control law $u = Kx$.

Now condition (4.4) holds for all x and u if

$$(A(t) + B(t)K)^T P + P(A(t) + B(t)K) + Q + K^T RK < 0$$

for all $t \geq 0$, which in turn is equivalent to

$$P^{-1}(A_i + B_iK)^T + (A_i + B_iK)P^{-1} + P^{-1}(Q + K^T RK)P^{-1} < 0, \quad i = 1, \ldots, 4.$$

With the change of variables $W = P^{-1}$ and $Y = KP^{-1}$, we get the matrix inequality

$$WA_i^T + A_iW + B_iY + Y^T B_i^T + WQW + Y^T RY < 0, \quad i = 1, \ldots, 4.$$

Thus the modified problem can be solved via the following LMI problem with $K = YW^{-1}$:

$$\min_{\gamma, W, Y} \gamma$$

subject to

$$\begin{bmatrix} WA_i^T + A_iW + BY + Y^T B^T & W & Y^T \\ W & -Q^{-1} & 0 \\ Y & 0 & -R^{-1} \end{bmatrix} < 0, \quad i = 1, \ldots, 4$$

and

$$\begin{bmatrix} \gamma & x_0^T \\ x_0 & W \end{bmatrix} > 0.$$

From here on, we consider this modified problem, and take γ as the performance index. We first illustrate a tradeoff between uncertainty bounds and the performance, and then show the conservativeness resulting from the standard worst-case design.

Suppose that we are given a parameter distribution as depicted in Figure 4.1 based on some a priori knowledge and/or measurement:

case 1: $0.08 \leq k(t) \leq 0.32, \quad 0.008 \leq f(t) \leq 0.032;$

case 2: $0.12 \leq k(t) \leq 0.28, \quad 0.012 \leq f(t) \leq 0.028.$

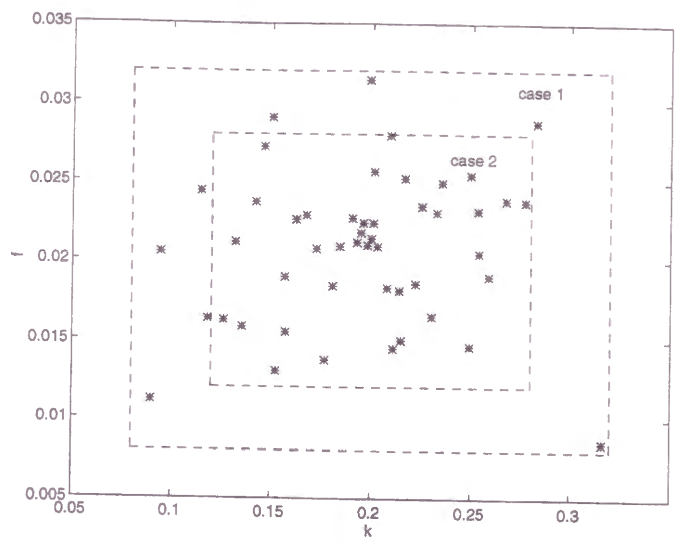

Figure 4.1: Parameter distribution

Case 1 corresponds to $\Omega(1)$ whereas case 2 represents $\Omega(2/3)$. Clearly $\Omega(2/3) \subset \Omega(1)$, and the standard robust control setup solves a design problem for case 1, since it covers all the parameter uncertainties.

We now solve the above LMI problem for both cases 1 and 2 with $Q = I$, $R = 1$, and obtain the matrices K_1 and K_2 that minimize γ's for $\Omega(1)$ and $\Omega(2/3)$, respectively. We then analyze the control performance for θ in the closed-loop systems obtained with K_1 and K_2.

$$\gamma_j(\theta) := \min_{\gamma, P} \gamma$$

subject to

$$\gamma - x_0^T P x_0 > 0, \qquad P > 0$$

and

$$(A_i(\theta) + BK_j)^T P + P(A_i(\theta) + BK_j) + Q + K_j^T RK_j < 0, \qquad i = 1, \ldots, 4,$$

where $\gamma_1(\theta)$ and $\gamma_2(\theta)$ are the above γ's according as the feedback gain is K_1 or K_2. Figure 4.2 shows the plots of $\gamma_1(\theta), \gamma_2(\theta)$. It shows the relation between the uncertain parameter ranges evaluated for synthesis and the resultant control performance. The relation exhibits a tradeoff between evaluated uncertainty ranges and control performance.

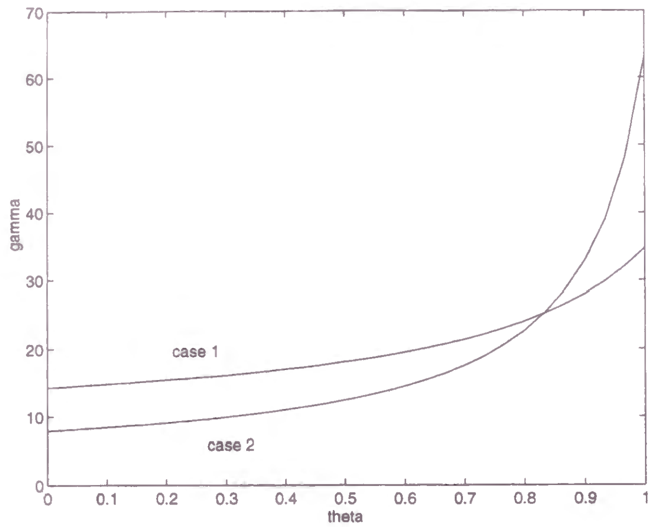

Figure 4.2: Tradeoff curves

Since the parameter distribution is known as $*$ in Figure 4.1, we see that the plant data appear more frequently for a smaller θ. On the other hand, $\gamma_2(\theta)$ shows better performance for a smaller θ, i.e., a smaller uncertainty range, in comparison with $\gamma_1(\theta)$. Hence K_1 obtained in standard robust control is more conservative from the viewpoint of the average performance, while K_2 is more adequate when taking the parameter distribution into account.

In view of such a tradeoff, one obviously wishes to find a compromise between performance and robustness.

4.2 Control System Design Considering the Tradeoff between Uncertainty and Performance

In this section, we first give a more general plant model such that the parameter distribution can be reflected on the control performance. We then give the problem formulation and discuss a property of tradeoff curves on control synthesis and analysis. Finally, we present an approximate method for this problem.

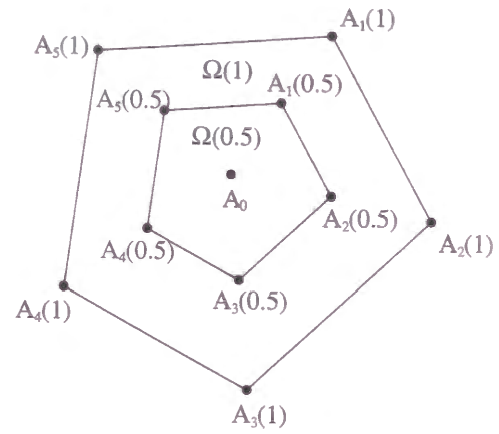

Figure 4.3: Matrix polytope $(L = 5)$

4.2.1 Models with Parameter Distribution

Consider the following linear system with time-varying parameters:

$$\begin{aligned} \frac{d}{dt}x(t) &= A(\xi)x(t) + Bu(t), \qquad x(0) = x_0, \\ &\quad A(\xi) \in \Omega(\theta), \end{aligned} \tag{4.5}$$

where $x \in \Re^{n_x}$ is the state of the plant, $u \in \Re^{n_u}$ is the control input and A is a matrix valued function of an uncertain random variable $\xi \in \Re^{n_\xi}$, and $\Omega(\theta)$ is a polytopic uncertainty set parametrized by $\theta \in [0,1]$ as follows:

$$\begin{aligned} \Omega(\theta) &:= \mathbf{Co}\{A_1(\theta), \ldots, A_L(\theta)\} \\ &\quad A_1(\theta) := A_0 + v_1(\theta)\hat{A}_1, \quad v_1(0) = 0, \\ &\quad \vdots \\ &\quad A_L(\theta) := A_0 + v_L(\theta)\hat{A}_L, \quad v_L(0) = 0, \\ &\quad A_0 \in \mathbf{Co}\{\hat{A}_1, \ldots, \hat{A}_L\}. \end{aligned}$$

Here $v_1(\theta), \ldots, v_L(\theta) : [0,1] \to [0,1]$ are monotone increasing functions. Figure 4.3 shows an example of $\Omega(\theta)$ where $\theta = 0, 0.5, 1$, and $L = 5$.

For the system introduced above, note

$$\theta_1 \le \theta_2 \Leftrightarrow \Omega(\theta_1) \subseteq \Omega(\theta_2).$$

Our objective is to find a controller such that the control performance is weighted with some kind of weighting functions that takes data distribution into account.

4.2.2 Control System Design Considering the Tradeoff

Let us first take the following performance index for the plant (4.5).

$$J(\theta) = \sup_{A(\cdot)\in\Omega(\theta)} \int_0^\infty (x(t)^T Q x(t) + u(t)^T R u(t))dt$$

where $Q > 0$, $R > 0$.

As already noted in the previous section, it is difficult to minimize this functional. We thus consider the following relaxed problem which gives an upper bound for $J(\theta)$.

$$\gamma_s(\theta) := \min_{\gamma,W,Y} \gamma \tag{4.6}$$

subject to

$$\begin{bmatrix} WA_i^T(\theta) + A_i(\theta)W + BY + Y^T B^T & W & Y^T \\ W & -Q^{-1} & 0 \\ Y & 0 & -R^{-1} \end{bmatrix} < 0, \qquad i = 1, \ldots, L$$

and

$$\begin{bmatrix} \gamma & x_0^T \\ x_0 & W \end{bmatrix} > 0.$$

The (performance) analysis counterpart (for each fixed K) is given by computing the following γ_K:

$$\gamma_K(\theta) := \min_{\gamma,P} \gamma \tag{4.7}$$

subject to

$$\gamma - x_0^T P x_0 > 0, \qquad P > 0$$

and

$$\begin{bmatrix} \begin{pmatrix} (A_i(\theta) + BK)^T P + \\ P(A_i(\theta) + BK) + Q \end{pmatrix} & K^T \\ K & -R^{-1} \end{bmatrix} < 0, \quad i = 1, \ldots, L.$$

These problems are a family of LMI problems parametrized by θ. If $Y(\theta)$ and $W(\theta)$ are the optimal solutions of the LMI problem (4.6) for a fixed θ, then $K(\theta) := Y(\theta)W(\theta)^{-1}$ gives the optimal gain matrix. Clearly there exists a tradeoff between the size of the evaluated uncertainty range θ and control performance $\gamma_s(\theta)$.

Definition 4.1 (U-P tradeoff curve) Define a *U-P (Uncertainty-Performance) tradeoff curve* to be the one given by the optimal value function $\gamma_s(\theta)$ for the plant (4.5) and the LMI problem (4.6). Also, define a *U-P controller* to be an optimal solution $K(\theta)$ on U-P tradeoff curves.

In the closed-loop system with controller K, a tradeoff between the size of uncertainty and the controller is provided by the optimal value function $\gamma_K(\theta)$ of the LMI problem (4.7).

Definition 4.2 (Guaranteed performance curve) Define a *guaranteed performance curve* to be the optimal value function $\gamma_K(\theta)$ for the plant (4.5), the control law $u = Kx$ and the LMI problem (4.7).

These curves $\gamma_s(\theta)$ and $\gamma_K(\theta)$ are both monotone increasing and piecewise smooth. But they are not convex in general. In connection with this, the following proposition holds.

Proposition 4.1 *For the U-P tradeoff curve $\gamma_s(\theta)$ and the guaranteed performance function $\gamma_{K(\theta_0)}(\theta)$ in the closed-loop system with U-P controller $K(\theta_0)$, the following relations hold:*

$$\gamma_{K(\theta_0)}(\theta) \geq \gamma_s(\theta), \qquad \gamma_{K(\theta_0)}(\theta_0) = \gamma_s(\theta_0).$$

Proof Define For given θ, γ, define the feasibility sets of the LMI problems (4.6) and (4.7) respectively as follows:

$$\begin{aligned} \mathcal{R}_s(\theta,\gamma) &:= \{(W,Y)|\ W \text{ and } Y \text{ satisfy (4.6)}\}, \\ \mathcal{R}_{K(\theta_0)}(\theta,\gamma) &:= \{(W,Y)|\ W = P^{-1}, Y = K(\theta_0)P^{-1}, \text{ and } P \text{ satisfies (4.7)}\}. \end{aligned}$$

Putting $W = P^{-1}$, $Y = KP^{-1}$ in (4.6), and using the Schur complement (e.g., [8]) yields the constraint (4.7). Since $K(\theta_0)$ is a constant matrix in $\mathcal{R}_{K(\theta_0)}(\theta,\gamma)$, $\mathcal{R}_{K(\theta_0)}(\theta,\gamma) \subseteq \mathcal{R}_s(\theta,\gamma)$ follows. This readily implies $\gamma_{K(\theta_0)}(\theta) \geq \gamma_s(\theta)$. On the other hand, if we regard

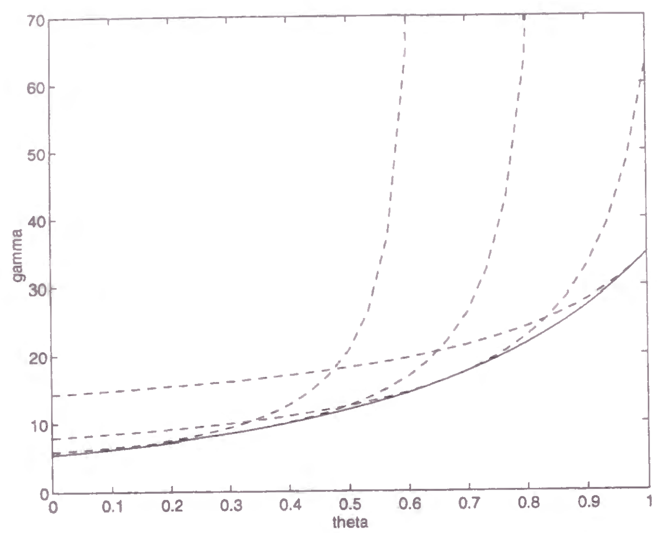

Figure 4.4: U-P Tradeoff curve and guaranteed performance curves

K as a free parameter, then the reverse procedure yields (4.6) from the constraint (4.7). This means that the optimal solution $Y(\theta_0)$, $W(\theta_0)$ and the optimum $\gamma_s(\theta_0)$ for (4.6) satisfies (4.7) with $P = W(\theta_0)^{-1}$, $K = Y(\theta_0)W(\theta_0)^{-1}$, $\gamma = \gamma_s(\theta_0)$. Hence the lower bound $\gamma_s(\theta)$ is attained at $\theta = \theta_0$. ■

Figure 4.4 shows the U-P tradeoff curve (solid) and the guaranteed performance curves $\gamma_{K(0)}$, $\gamma_{K(1/3)}$, $\gamma_{K(2/3)}$ and $\gamma_{K(1)}$ (dashed). Note also that since Q is assumed to be positive definite, the closed-loop stability is guaranteed for every θ if the guaranteed performance curve remains to be finite in the overall region.

Clearly the U-P tradeoff curve $\gamma_s(\theta)$ gives the lowest bound for the relaxed performance, and at any specific point θ_0, $\gamma_{K(\theta_0)}$ agrees with $\gamma_s(\theta_0)$. However, at other points there is a gap between the two. If we could observe θ and use $K(\theta)$ adaptively, then we would be able to attain the best performance $\gamma_s(\theta)$. This cannot be achieved by a single controller, but it is still desirable to reduce the performance gap in the over all range of θ:

Problem 1: Find a single control law $u = Kx$ that gives rise to an acceptable guaranteed performance curve $\gamma_K(\theta)$ for system (4.5).

To this end, we introduce a weighting function $w(\theta) : [0,1] \to \Re_+$ for $\gamma_K(\theta)$. The idea here is that we can make a compromise in the performance curve, by placing more

weights on those θ's where the above gaps are to be small. On the other hand, one can place less emphasis on those θ's that are of less interest to us. Such a weight is thus a design parameter, and in this chapter we suppose that it enters into the objective function $f(\gamma_K, w; \theta)$ either additively or multiplicatively as follows:

- Additive weighting function

$$f(\gamma_K, w; \theta) = \gamma_K(\theta) + w(\theta).$$

- Multiplicative weighting function

$$f(\gamma_K, w; \theta) = \gamma_K(\theta)w(\theta).$$

One way of dealing with this situation is to employ the probability density distribution $\mu(\theta)$ that governs the probability for $A(\xi)$ to belong to the set $\Omega(\theta)$:

$$f(\gamma_K, p_\theta; \theta) = \gamma_K(\theta)\mu(\theta).$$

Of course, $\mu(\theta)$ cannot be known exactly, but it may be approximated by some weight function derived from the empirical data when we try to identify the plant parameters.

Problem 1 is a minimization problem of a weighted objective function $f(\gamma_K, w; \theta)$ subject to LMI constraints (4.7) for all $\theta \in [0,1]$. Thus a solution of Problem 1 is given by a solution K of the following min-max problem.

$$\min_{\gamma_K(\theta), K, P(\theta)} \max_{\theta \in [0,1]} f(\gamma_K, w; \theta)$$

subject to

$$\gamma_K(\theta) - x_0^T P(\theta) x_0 > 0, \qquad P(\theta) > 0$$

and

$$\begin{bmatrix} \begin{pmatrix} (A_i(\theta) + BK)^T P(\theta) + \\ P(\theta)(A_i(\theta) + BK) + Q \end{pmatrix} & K^T \\ K & -R^{-1} \end{bmatrix} < 0, \quad i = 1, \ldots, L,$$

where $\gamma_K(\theta)$ and $P(\theta)$ are variables depending on θ. Note that K does not depend on θ. This min-max problem is hard to tract, but it can be reduced to the following equivalent minimization problem:

$$\min_{\nu, \gamma_K(\theta), K, P(\theta)} \nu \tag{4.8}$$

subject to

$$\nu - f(\gamma_K, w; \theta) \geq 0, \qquad \gamma_K(\theta) - x_0^T P(\theta) x_0 > 0, \qquad P(\theta) > 0$$

and

$$\begin{bmatrix} \begin{pmatrix} (A_i(\theta) + BK)^T P(\theta) + \\ P(\theta)(A_i(\theta) + BK) + Q \end{pmatrix} & K^T \\ K & -R^{-1} \end{bmatrix} < 0, \quad i = 1, \ldots, L, \quad \text{for all } \theta \in [0, 1].$$

Unfortunately this semi-infinite programming problem (i.e., problem with an infinite number of constraints) is a non-convex optimization problem and is not easily solvable. We will present an approximate solution for this in the next section.

Figure 4.5 gives a geometrical interpretation of the problem (4.8). The solid curve shows the guaranteed performance curve $\gamma_K(\theta)$ for a feasible solution K. The shaded region shows the set

$$\begin{aligned} \mathcal{F}_K \ := \ \{(\theta, \gamma) \mid \ & \gamma - x_0^T P(\theta) x_0^T > 0, \quad P(\theta) > 0, \\ & \begin{bmatrix} \begin{pmatrix} (A_i(\theta) + BK)^T P(\theta) + \\ P(\theta)(A_i(\theta) + BK) + Q \end{pmatrix} & K^T \\ K & -R^{-1} \end{bmatrix} < 0, \quad i = 1, \ldots, L \ \} \end{aligned}$$

for K and the dashed curve shows the boundary of the level set

$$\mathcal{S}_\nu := \{(\theta, \gamma) \mid \nu - f(\gamma, w; \theta) \geq 0 \ \}.$$

The optimization problem (4.8) is to minimize ν such that the intersection of $\mathcal{S}_\nu$ and $\mathcal{F}_K$ is not empty for all $\theta \in [0, 1]$:

$$\min_{\nu, K, P(\theta)} \{\nu \mid \mathcal{F}_K \cap \mathcal{S}_\nu \neq \emptyset, \text{ for all } \theta \in [0, 1]\}. \tag{4.9}$$

If ν remains constant even when a larger value is assumed at $w(\theta)$ (to give more weight at θ), the supremum of γ for $\mathcal{S}_\nu$ becomes smaller. The shape of the boundary of $\mathcal{S}_\nu$ thus depends on the weight $w(\theta)$. Hence problem (4.9) may be interpreted as that of finding K that gives $\gamma_K(\theta)$ (i.e., the boundary of $\mathcal{F}_K$) as close as possible to the desired shape of the boundary of $\mathcal{S}_\nu$ determined by $w(\theta)$.

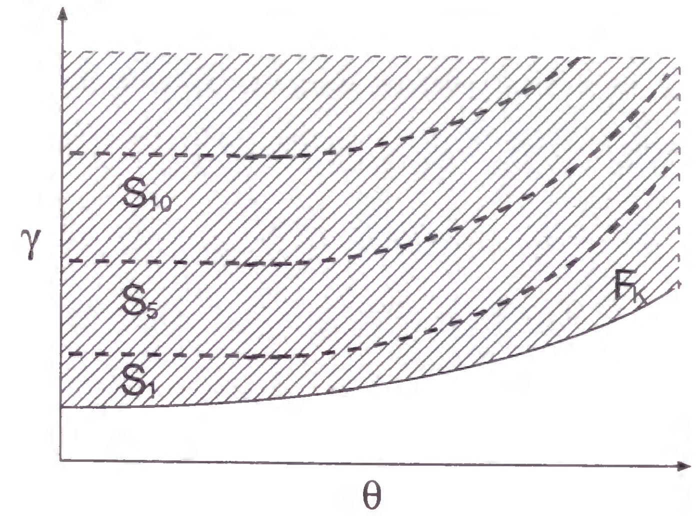

Figure 4.5: Geometrical interpretation of (4.8)

4.2.3 Approximate Method

To remedy the difficulty in solving (4.8), we replace the constraints by those induced by picking only a finite number of points for θ. Take $\theta_1, \ldots, \theta_M$ from $\theta \in [0, 1]$ so that only M performance levels of uncertainty can be taken into account. The problem (4.8) is then reduced to the following problem:

$$\min_{\nu, \gamma_j, K, P_j} \nu \tag{4.10}$$

subject to

$$\nu - f(\gamma_K, w; \theta_j) \geq 0, \quad \gamma_j - x_0^T P_j x_0 > 0, \quad P_j > 0,$$

$$\begin{bmatrix} \begin{pmatrix} (A_{ij} + BK)^T P_j + \\ P_j(A_{ij} + BK) + Q \end{pmatrix} & K^T \\ K & -R^{-1} \end{bmatrix} < 0, \qquad i = 1, \ldots, L, \qquad j = 1, \ldots, M,$$

where $A_{ij} = A_i(\theta_j)$, $P_j = P(\theta_j)$, $\gamma_j = \gamma_K(\theta_j)$ and $f(\gamma_K, w; \theta_j)$ is a linear function of γ_j. This is an optimization problem with a finite number of BMI constraints.

In doing this, one needs to guarantee that this procedure can converge to the optimal solution when we increase the number of points θ_j. While this may not give a practical bound, we can at least answer this question in the following form:

Theorem 4.1 *Fix a sufficiently large $M > 0$ and we suppose that search of K occurs in the ball $B_M := \{K : \|K\| \le M\}$. Suppose that the semi-infinite problem (4.8) admits an optimal value ν_{opt}, where controllers K are to be found in B_M. Suppose also that the function f appearing in (4.8) or (4.10) is continuous in its variables. Then*

$$\nu_{opt} = \sup_J \nu_J \tag{4.11}$$

where the right-hand side denotes the supremum of all ν_J's that are a solution to (4.10) over all finite sets J of points of θ in $[0, 1]$.

In other words, the true optimal value may be obtained as a limit of the above finite-point approximations.

Proof Let us first fix a small $\delta > 0$ and consider the BMI constraints in (4.8) and (4.10) with margin δ, i.e., we consider $\nu - f(\gamma_K, w; \theta_j) \ge \delta$,

$$\begin{bmatrix} \begin{pmatrix} (A_{ij} + BK)^T P_j + \\ P_j(A_{ij} + BK) + Q \end{pmatrix} & K^T \\ K & -R^{-1} \end{bmatrix} \le -\delta I$$

and $\gamma_j - x_0^T P_j x_0 \ge \delta$ in place of those in (4.10) etc. By suitably changing the signs, we represent these constraints as $\Gamma(\Theta_N, K, \nu) \le -\delta$ where Θ_N denotes the finite set $\{\theta_1, \ldots, \theta_N\}$. For the constraint corresponding to (4.8) on the over all range in θ, we denote it by $\Gamma(K, \nu) \le -\delta$ since it does not depend on θ. If we need to refer to such a constraint at each point θ, we will denote it by $\Gamma(\theta, K, \nu)$ keeping its dependence on θ by the lower case letter. Needless to say, $\Gamma(K, \nu) := \sup_{0 \le \theta \le 1} \Gamma(\theta, K, \nu)$.

We first consider the convergence property with this margin δ taken into account. Note first that since the original problem (corresponding to $\delta = 0$) is assumed to be feasible for $\nu = \nu_{opt}$, there exists, for any $\nu > \nu_{opt}$, a K with $\|K\| \le M$ such that $\Gamma(\theta, K, \nu) < 0$ for all θ. Since this constraint is a continuous function in θ, this inequality actually implies $\Gamma(\theta, K, \nu) \le -\delta < 0$ for some $\delta > 0$. Thus the modified problem is solvable for at least sufficiently small δ.

We prove equality (4.11) for this case first. It is clear that

$$\nu_{opt} \ge \sup_J \nu_J$$

because each term inside the supremum on the right means less constraint than that on the left. We need to prove the reverse inequality, i.e., the left-hand side can be attained as a limit of terms on the right. Take a sequence of finite sets Θ_N such that

- $\Theta_N \subset \Theta_{N+1}$ for all N, i.e., Θ_{N+1} is a refinement of Θ_N;
- $|\Theta_N| \to 0$, i.e., $\max_{1 \le i \le N-1} |\theta_{i+1} - \theta_i| \to 0$ as $N \to \infty$.

The union of all such Θ_N gives a dense subset of $[0, 1]$. Let $\{\nu_N\}$ be the corresponding optimal value under constraint $\Gamma(\Theta_N, K, \nu_N) \le -\delta$. Without loss of generality, we may assume that $\{\nu_N\}$ is monotone increasing (non-decreasing). Then for each N there exists K_N such that $\Gamma(\Theta_N, K_N, \nu_N) \le -\delta$. Since the ball B_M of matrices is compact, there exists a subsequence of K_N that is convergent to K in B_M. Let $\nu := \sup \nu_N$. Then by continuity, we have $\Gamma(K, \nu) \le -\delta$. This means that the right-hand side is not less than the left-hand side in (4.11). This implies the desired equality for this case.

The optimum for the original problem with $\delta = 0$ is obtained as the infimum of the above cases for $\delta > 0$, although in the limiting case the existence of an optimal controller K is no longer guaranteed. In any event, by the continuity of the constraints, the equality (4.11) for $\delta > 0$ carries over to the limit, and this completes the proof. ■

While the theorem above guarantees the convergence of the approximation, a practical and global method for a general BMI problem applicable to the present situation is unfortunately not known at present. By solving LMI problems iteratively, however, a suboptimal solution can be obtained. In the next section, we obtain a solution from U-P controllers by a line search on $\theta \in [0, 1]$.

4.3 Numerical Example

We take the numerical example in Section 4.1 again. Suppose that the distribution of uncertain parameters is given by some measurement as in Figure 4.2 and that the plant (4.1) is obtained from the density function $\mu(\theta)$. Considering the distribution of uncertain parameters, we set a multiplicative weighting function:

$$w(\theta) := \mu(\theta) = (7 - 6\theta)/4$$

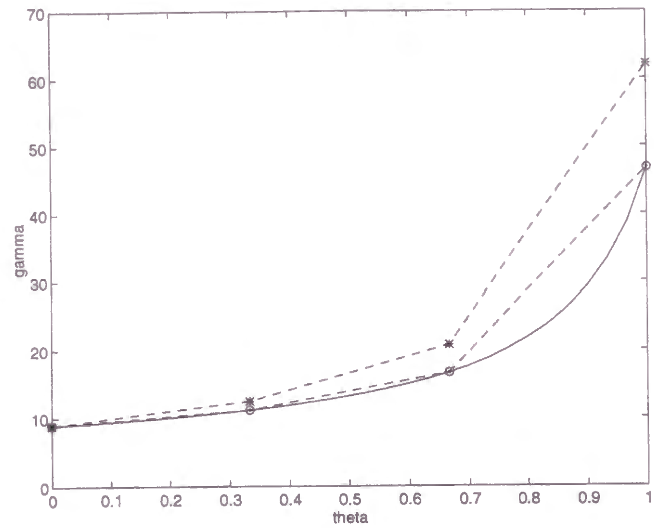

Figure 4.6: $\gamma_{K^*}(\theta)$, $\partial\mathcal{F}_{K^*}$ and $\partial\mathcal{S}_{\nu^*}$

and select performance levels:

$$\theta_1 = 0, \theta_2 = 1/3, \theta_3 = 2/3, \theta_4 = 1.$$

The following results is obtained by the method of the previous section:

$$K^* = [-6.10 \quad 4.07 \quad -3.78 \quad -1.08], \quad \nu^* = 62.1, \quad \theta^* = 0.75.$$

The approximate guaranteed performance is 13.3 in this case while that in the case of $K(1)$ it is 17.8. We see that the proposed method is better from the viewpoint of the average performance. Figure 4.6 shows the guaranteed performance curve $\gamma_{K^*}(\theta)$ (solid), the points ($\circ$) of the boundary of $\mathcal{F}_{K^*}$ and the points ($*$) on the boundary of $\mathcal{S}_{\nu^*}$.

4.4 Extensions to the H^∞ Case

Although the presented method has been devoted to the LQR design framework, it can be easily extended to the design method based on other performance indices. In the H^∞ robust control design, we consider the following system:

$$\begin{aligned}
\frac{d}{dt}x(t) &= A(\xi)x(t) + B_1 w(t) + B_2 u(t) \\
z &= C_1 x(t) + D_{11} w(t) + D_{12} u(t) \\
& A(\xi) \in \Omega(\theta),
\end{aligned}$$

where $w \in \Re^{n_w}$ is the exogenous input, $z \in \Re^{n_z}$ is the control output, and x, u, ξ and Ω are defined in the same manner as the LQR case. Also, the l^2 induced norm is used as the H^∞ performance index. For the above system and performance index, the synthesis problem corresponding to (4.6) is as follows:

$$\begin{aligned}
\hat{\gamma}_{\infty,s}(\theta) &:= \sqrt{\gamma_{\infty,s}(\theta)}, \\
\gamma_{\infty,s}(\theta) &:= \min_{\gamma,W,Y} \quad \gamma
\end{aligned}$$

subject to

$$\begin{bmatrix} WA_i^T(\theta) + A_i(\theta)W + B_2Y + Y^TB_2^T & B_1 & WC_1^T + Y^TD_{12}^T \\ B_1^T & -I & D_{11}^T \\ C_1W + D_{12}Y & D_{11} & -\gamma I \end{bmatrix} < 0, \qquad i = 1, \ldots, L,$$

and

$$W > 0.$$

The analysis problem corresponding to (4.7) is as follows:

$$\begin{aligned}
\hat{\gamma}_{\infty,K}(\theta) &:= \sqrt{\gamma_{\infty,K}(\theta)}, \\
\gamma_{\infty,K}(\theta) &:= \min_{\gamma,P} \quad \gamma
\end{aligned}$$

subject to

$$\begin{bmatrix} (A_i(\theta) + B_2K)P + P(A_i(\theta) + B_2K) + Q & B_1 & P(C_1 + D_{12}K)^T \\ B_1^T & -I & D_{11}^T \\ (C_1 + D_{12}K)P & D_{11} & -\gamma I \end{bmatrix} < 0, \quad i = 1, \ldots, L,$$

and

$$P > 0.$$

The rest of the development is similar to that given in the previous sections.

Chapter 5

LMI-based Model Predictive Control with Rate Constraints

Model predictive control (MPC) has become an increasingly popular control design method in the process industries. Although there exist many different MPC techniques, they are all based on the same concept: They involve the solution of an on-line optimization problem subject to various constraints in order to determine optimal future control inputs. Usually, only the first control action is implemented and at the next sampling time, system measurements are used to update the optimization problem. Such a computation is no longer a burden with today's advanced computing power. However, the fundamental issue in such a scheme is that it is generally difficult to guarantee robust stability and performance.

Kothare et al. [35] proposed a robust MPC method based on LMIs. The method allows consideration of amplitude constraints on the control input and output and guarantees stability for uncertain systems. Moreover, the method has the ability to deal with a variety of constraints in the framework of the state-feedback synthesis [8].

On the other hand, the MPC method does not deal with such rate constraints. In general, actuators have a limited rate as well as a limited range of action, as is the case of control valves which are limited by a maximum slew rate as well as a fully closed and fully open position. In practice, both amplitude and rate of the control input are usually taken into account in several existing MPC techniques [22, 10, 9]. In some cases, the rate of the control output is also controlled by such reasons as constructive safety [9, 7].

In this chapter, we present LMI conditions for the rate limits of the control input and output in the framework of the robust LMI-based MPC. A numerical example illustrates that the MPC method with these rate constraints provides better performance. The results of this chapter are based on [60].

5.1 Robust LMI-Based MPC

Kothare et al. [35] deal with two types of system models for robust control, i.e, polytopic systems and linear systems with structured feedback uncertainty. In this chapter, we focus on only polytopic system models, but the results for linear system models with structured feedback uncertainty can be derived in a similar manner [60].

We will state the outline of robust LMI-based MPC [35]. Consider the following linear system:

$$\begin{aligned} x(k+1) &= A(k)x(k) + B(k)u(k) \\ y(k) &= Cx(k) \qquad\qquad (5.1) \\ [A(k)\ B(k)] &\in \Omega \\ \Omega &= \mathbf{Co}\{[A_1\ B_1], [A_2\ B_2], \ldots, [A_L\ B_L]\}, \qquad (5.2) \end{aligned}$$

where $u(k) \in \Re^{n_u}$ is the control input, $x(k) \in \Re^{n_x}$ is the state of the plant, $y(k) \in \Re^{n_y}$ is the control output, and the set Ω is the polytope.

In the MPC framework, the robust performance objective to be minimized at each sampling time k is as follows:

$$\min_{u(k+i|k), i\geq 0} \ \max_{[A(k+i)\ B(k+i)]\in\Omega, i\geq 0} J(k), \tag{5.3}$$

where

$$J(k) = \sum_{i=0}^{\infty} \left(x(k+i|k)^T Q_1 x(k+i|k) + u(k+i|k)^T R u(k+i|k) \right)$$

and Q_1, R are symmetric weighting matrices; $Q_1 \geq 0$, $R > 0$. Here

$u(k|k)$: control action to be implemented at time k;

$x(k|k)$: state measured at time k;

$y(k|k)$: output measured at time k;

$u(k+i|k)$: control action at time $k+i$; computed by the optimization problem (5.3) or the SDP in Theorem 5.1 at time k;

$x(k+i|k)$: state at time $k+i$, predicted based on measurements at time k;

$y(k+i|k)$: output measured at time $k+i$, predicted based on measurements at time k.

The problem (5.3) is a min-max problem. The maximization is over the set Ω and corresponds to choosing that time-varying plant $[A(k+i)\ B(k+i)] \in \Omega, i \geq 0$ which, if used as a model for predictions, would lead to the largest or worst-case value of $J(k)$ among all plants in Ω. This worst-case value is minimized over present and future control actions $u(k+i|k), i = 0, 1, \ldots, m$. This min-max problem, though convex for finite m, is not computationally tractable, and as such has not been addressed in the MPC literature.

We address problem (5.3) by first deriving an upper bound on the robust performance objective. We then minimize this upper bound with a constant state-feedback control law $u(k+i|k) = Fx(k+i|k), i \geq 0$.

Consider a quadratic function $V(x) = x^T P x, P > 0$ of the state $x(k|k)$ of the system (5.1) with $V(0) = 0$. At sampling time k, suppose V satisfies the following inequality for all $x(k+i|k), u(k+i|k), i \geq 0$ satisfying (5.1), and for any $[A(k+i)\ B(k+i)] \in \Omega, i \geq 0$:

$$\begin{aligned} & V(x(k+i+1|k)) - V(x(k+i|k)) \\ & \leq -\left(x(k+i|k)^T Q_1 x(k+i|k) + u(k+i|k)^T R u(k+i|k) \right). \end{aligned} \tag{5.4}$$

For the robust performance objective function to be finite, we must have $x(\infty|k) = 0$ and hence, $V(x(\infty|0)) = 0$. Summing (5.4) from $i = 0$ to $i = \infty$, we get

$$-V(x(k|k)) \leq -J_\infty(k).$$

Thus

$$\max_{[A(k+i)\ B(k+i)]\in\Omega, i\geq 0} J_\infty(k) \leq V(x(k|k)). \tag{5.5}$$

This gives an upper bound for the robust performance objective. Thus the goal of the robust MPC algorithm has been redefined to synthesize, at each time step k, a constant state-feedback control law $u(k+i|k) = Fx(k+i|k)$ to minimize this upper bound $V(x(k|k))$. As is standard in MPC, only the first computed input $u(k|k) = Fx(k|k)$ is implemented. At the next sampling time, the state $x(k+1|k+1)$ is measured and the optimization is repeated to recompute F. The following theorem gives us conditions for the existence of the appropriate $P > 0$ satisfying (5.4) and the corresponding state-feedback matrix F [35].

Theorem 5.1 *Let $x(k) = x(k|k)$ be the state of the uncertain system (5.1) measured at sampling time k. Assume that there are no constraints on the control input and output. The state-feedback matrix F in the control law $u(k+i|k) = Fx(k+i|k), i \geq 0$ which minimizes the upper bound on the robust performance objective function at sampling time k is given by*

$$F = YQ^{-1} \tag{5.6}$$

where $Q > 0$ and Y are obtained from the solution (if it exists) to the following SDP:

$$\min_{\gamma,Q,Y} \gamma \tag{5.7}$$

subject to

$$\begin{bmatrix} 1 & x(k|k)^T \\ x(k|k) & Q \end{bmatrix} \geq 0, \tag{5.8}$$

$$\begin{bmatrix} Q & QA_j^T + Y^T B_j^T & QQ_1^{\frac{1}{2}} & Y^T R^{\frac{1}{2}} \\ A_jQ + B_jY & Q & 0 & 0 \\ Q_1^{\frac{1}{2}}Q & 0 & \gamma I & 0 \\ R^{\frac{1}{2}}Y & 0 & 0 & \gamma I \end{bmatrix} \geq 0, \quad j = 1, 2, \ldots, L. \tag{5.9}$$

Proof Minimization of $V(x(k|k)) = x(k|k)^T Px(k|k)$, $P > 0$ is equivalent to

$$\min_{\gamma,P} \{\gamma \mid x(k|k)^T Px(k|k) \leq \gamma\}.$$

Defining $Q = \gamma P^{-1} > 0$ and using Schur complements, this is equivalent to

$$\min_{\gamma,Q} \gamma$$

subject to

$$\begin{bmatrix} 1 & x(k|k)^T \\ x(k|k) & Q \end{bmatrix} \geq 0,$$

which establishes (5.7) and (5.8). It remains to prove (5.6) and (5.9). The quadratic function V is required to satisfy (5.4). Substituting $u(k+i|k) = Fx(k+i|k), i \geq 0$ and the state space (5.1), inequality (5.4) becomes:

$$x(k+i|k)^T \Big((A(k+i) + B(k+i)F)^T P(A(k+i) + B(k+i)F) - P$$
$$F^T RF + Q_1 \Big) x(k+i|k) \leq 0.$$

This is satisfied for all $i \geq 0$ if

$$(A(k+i) + B(k+i)F)^T P(A(k+i) + B(k+i)F) - P + F^T RF + Q_1 \leq 0. \tag{5.10}$$

Substituting $P = \gamma Q^{-1}$, $Q > 0$, $Y = FQ$, pre- and post-multiplying by Q (which leaves the inequality unaffected), and using Schur complements, we see that this is equivalent to

$$\begin{bmatrix} Q & QA(k+i)^T + Y^T B(k+i)^T & QQ_1^{\frac{1}{2}} & Y^T R^{\frac{1}{2}} \\ A(k+i)Q + B(k+i)Y & Q & 0 & 0 \\ Q_1^{\frac{1}{2}}Q & 0 & \gamma I & 0 \\ R^{\frac{1}{2}}Y & 0 & 0 & \gamma I \end{bmatrix} \geq 0. \tag{5.11}$$

Inequality (5.11) is affine in $[A(k+i)\, B(k+i)]$. Hence it is satisfied for all

$$[A(k+i)\ B(k+i)] \in \Omega = \mathbf{Co}\{[A_1\ B_1], [A_2\ B_2], \ldots, [A_L\ B_L]\}$$

if and only if there exist $Q > 0$, $Y = FQ$ and γ such that

$$\begin{bmatrix} Q & QA_j^T + Y^T B_j^T & QQ_1^{\frac{1}{2}} & Y^T R^{\frac{1}{2}} \\ A_jQ + B_jY & Q & 0 & 0 \\ Q_1^{\frac{1}{2}}Q & 0 & \gamma I & 0 \\ R^{\frac{1}{2}}Y & 0 & 0 & \gamma I \end{bmatrix} \geq 0, \quad j = 1, 2, \ldots, L.$$

The feedback matrix is then given by $F = YQ^{-1}$. This established (5.6) and (5.9). ■

Remark 5.1 Strictly speaking, the variables in the above optimization should be denoted by Q_k, F_k, Y_k etc. to emphasize that they are computed at time k. For notational convenience, we omit the subscript. We will, however, briefly utilize this notation in the robust stability proof. Closed loop stability of the receding horizon state-feedback control law given in Theorem 5.1 will be established in Section 5.3

We show the lemma associated with the state of the system, which is used in the proofs of sufficient LMI constraints [35, 8].

Lemma 5.1 *Consider the system (5.1). At sampling time k, suppose there exist $Q > 0$, γ and $Y(= FQ)$ such that (5.9) holds. Also, suppose that $u(k+i|k) = Fx(k+i|k), i \geq 0$. Then if*

$$x(k|k)^T Q^{-1} x(k|k) \leq 1,$$

or equivalently,

$$x(k|k)^T P x(k|k) \leq \gamma \text{ with } P = \gamma Q^{-1},$$

then

$$\max_{[A(k+j)\ B(k+j)]\in\Omega, j\geq 0} x(k+i|k)^T Q^{-1} x(k+i|k) < 1, \quad i \geq 1, \tag{5.12}$$

or equivalently,

$$\max_{[A(k+j)\ B(k+j)]\in\Omega, j\geq 0} x(k+i|k)^T P x(k+i|k) < \gamma, \quad i \geq 1.$$

Thus $\mathcal{E} = \left\{z \mid z^T Q z \leq 1\right\} = \left\{z \mid z^T P z \leq \gamma\right\}$ is an invariant ellipsoid for the predicted states of the uncertain system.

Proof From the proof of Theorem 5.1, we know that (5.9) and (5.10) are equivalent and that (5.10) implies (5.4). Thus

$$\begin{aligned} & x(k+i+1|k)^T P x(k+i+1|k) - x(k+i|k)^T P x(k+i|k) \\ \leq\ & -x(k+i|k)^T Q_1 x(k+i|k) - u(k+i|k)^T R u(k+i|k) \\ <\ & 0 \qquad \text{since } Q_1 > 0. \end{aligned}$$

Therefore,

$$x(k+i+1|k)^T P x(k+i+1|k) < x(k+i|k)^T P x(k+i|k), \quad i \geq 0, \quad (x(k+i|k) \neq 0). \tag{5.13}$$

Thus, if $x(k|k)^T P x(k|k) \leq \gamma$, then $x(k+1|k)^T P x(k+1|k) < \gamma$. This argument can be continued for $x(k+2|k), x(k+3|k), \ldots$ and this completes the proof. ■

5.1.1 Amplitude Constraints

In this subsection, we show sufficient LMI constraints on the amplitude limits of the input [35, 8]. We consider Euclidean norm (l^2 norm) bounds and component-wise peak (l^∞ norm) bounds on the amplitude of the control input and output. Here, for a vector x, define

$$\|x\|_2^2 := x^T x, \quad \|x\|_\infty := \max_i |x_i|.$$

Lemma 5.2 (Amplitude constraint on the input (l^2 norm case)) *For the system (5.1), if there exist $Q > 0$ and Y satisfying (5.8), (5.9) and*

$$\begin{bmatrix} \alpha_1 I & Y \\ Y^T & Q \end{bmatrix} \geq 0, \tag{5.14}$$

then

$$\max_{i\geq 0} \|u(k+i|k)\|_2^2 \leq \alpha_1.$$

Lemma 5.3 (Amplitude constraint on the input (l^∞ norm case)) *For the system (5.1), if there exist $Q > 0$ and Y satisfying (5.8), (5.9) and*

$$\begin{bmatrix} X & Y \\ Y^T & Q \end{bmatrix} \geq 0 \quad \text{with } X_{jj} \leq \alpha_2, \quad j = 1, 2, \ldots, n_u, \tag{5.15}$$

then

$$\max_{i\geq 0} \|u(k+i|k)\|_\infty^2 \leq \alpha_2.$$

5.2 Rate Constraints on the Input and Output

Many MPC techniques usually take into account constraints such as amplitude and rate limits and have therefore been very successful in industry. In [35, 8], however, no LMI conditions for rate limits have been shown. In this section, we present LMI conditions for the rate limits of the input and output.

5.2.1 Rate Constraints on the Input

We first present LMI constraints on the rate limits of the input. In the case of l^2 norm, we obtain the following lemma.

Lemma 5.4 (Rate constraint on the input (l^2 norm case)) *For the system (5.1), if there exist $Q>0$ and Y satisfying (5.8), (5.9), (5.14) and*

$$\begin{bmatrix} \beta_1 Q & QA_j^T + Y^T B_j^T - Q \\ A_j Q + B_j Y - Q & \frac{1}{\alpha_1} Q \end{bmatrix} \geq 0, \quad j = 1, 2, \ldots, L, \tag{5.16}$$

then

$$\max_{i \geq 0} \|u(k+i+1|k) - u(k+i|k)\|_2^2 \leq \beta_1.$$

Proof Using Lemma 5.1 and $Q^{-1}Y^TYQ^{-1} \leq \alpha_1 Q^{-1}$ from (5.14), the following relation holds:

$$\begin{aligned}
& \max_{i \geq 0} \|u(k+i+1|k) - u(k+i|k)\|_2^2 \\
= & \max_{i \geq 0} \|F(A(k+i) + B(k+i)F - I)x(k+i|k)\|_2^2 \\
\leq & \max_{z \in \mathcal{E}, i \geq 0} \|YQ^{-1}(A(k+i) + B(k+i)YQ^{-1} - I)z\|_2^2 \\
= & \max_{\xi^T\xi \leq 1, i \geq 0} \|YQ^{-1}(A(k+i) + B(k+i)YQ^{-1} - I)Q^{1/2}\xi\|_2^2 \\
= & \max_{i \geq 0} \sigma_{\max}\left(YQ^{-1}(A(k+i) + B(k+i)YQ^{-1} - I)Q^{1/2}\right)^2 \\
= & \max_{i \geq 0} \lambda_{\max}\left(Q^{1/2}\left(A^T(k+i) + Q^{-1}Y^TB^T(k+i) - I\right)Q^{-1}Y^TYQ^{-1}\right. \\
& \qquad \left. \cdot \left(A(k+i) + B(k+i)YQ^{-1} - I\right)Q^{1/2}\right) \\
\leq & \max_{i \geq 0} \lambda_{\max}\left(Q^{-1/2}\left(QA^T(k+i) + Y^TB^T(k+i) - Q\right)\alpha_1 Q^{-1}\right. \\
& \qquad \left. \cdot (A(k+i)Q + B(k+i)Y - Q)\,Q^{-1/2}\right).
\end{aligned}$$

Therefore, if

$$\begin{aligned}
& Q^{-1/2}\left(QA^T(k+i) + Y^TB^T(k+i) - Q\right)\alpha_1 Q^{-1} \\
& \qquad \cdot (A(k+i)Q + B(k+i)Y - Q)\,Q^{-1/2} \leq \beta_1 I, \quad \text{for all } i \geq 0,
\end{aligned} \tag{5.17}$$

then $\max_{i \geq 0} \|u(k+i+1|k) - u(k+i|k)\|_2^2 \leq \beta_1$ is satisfied. Since $Q>0$, (5.17) is equivalent to

$$\begin{aligned}
& \left(QA^T(k+i) + Y^TB^T(k+i) - Q\right)\alpha_1 Q^{-1} \\
& \qquad \cdot (A(k+i)Q + B(k+i)Y - Q) \leq \beta_1 Q, \quad \text{for all } i \geq 0.
\end{aligned}$$

Furthermore, using Schur complements [8], we can see that (5.17) is equivalent to

$$\begin{bmatrix} \beta_1 Q & QA^T(k+i) + Y^TB^T(k+i) - Q \\ A(k+i)Q + B(k+i)Y - Q & \frac{1}{\alpha_1} \end{bmatrix} \geq 0, \quad \text{for all } i \geq 0.$$

Since the last inequality is affine in $[A(k+i)\ B(k+i)]$, it is satisfied for all

$$[A(k+i)\ B(k+i)] \in \Omega = \mathbf{Co}\{[A_1\ B_1], [A_2\ B_2], \ldots, [A_L\ B_L]\}$$

if and only if

$$\begin{bmatrix} \beta_1 Q & QA_j^T + Y^T B_j^T - Q \\ A_j Q + B_j Y - Q & \frac{1}{\alpha_1} Q \end{bmatrix} \geq 0, \quad j = 1, 2, \ldots, L$$

■

In the case of l^∞ norm, we obtain the following lemma.

Lemma 5.5 (Rate constraint on the input (l^∞ norm case)) *For the system (5.1), if there exist $Q>0$ and Y satisfying (5.8), (5.9), (5.15) and*

$$\begin{bmatrix} \beta_2 Q & A_j Q + B_j Y - Q \\ QA_j^T + Y^T B_j^T - Q & \frac{1}{\alpha_2} Q \end{bmatrix} \geq 0, \quad j = 1, 2, \ldots, L,$$

then

$$\max_{i \geq 0} \|u(k+i+1|k) - u(k+i|k)\|_\infty^2 \leq \beta_2.$$

Proof Using Lemma 5.1, the following relation holds:

$$\begin{aligned}
& \max_{i \geq 0} \|u(k+i+1|k) - u(k+i|k)\|_\infty^2 \\
= & \max_{i \geq 0} \|F\,(A(k+i) + B(k+i)F - I)\,x(k+i|k)\|_\infty^2 \\
\leq & \max_{z \in \mathcal{E}, i \geq 0} \|YQ^{-1}\left(A(k+i) + B(k+i)YQ^{-1} - I\right)z\|_\infty^2 \\
= & \max_{\xi^T\xi \leq 1, i \geq 0} \|YQ^{-1}\left(A(k+i) + B(k+i)YQ^{-1} - I\right)Q^{1/2}\xi\|_\infty^2 \\
\leq & \max_{i \geq 0, j} \left[YQ^{-1}\left(A(k+i) + B(k+i)YQ^{-1} - I\right)Q^{1/2}\right. \\
& \qquad \left. Q^{1/2}\left(A^T(k+i) + Q^{-1}Y^TB^T(k+i) - I\right)Q^{-1}Y^T\right]_{jj} \\
= & \max_{i \geq 0, j} \left[YQ^{-1}\,(A(k+i)Q + B(k+i)Y - Q)\,Q^{-1}\right. \\
& \qquad \left. \left(QA^T(k+i) + Y^TB^T(k+i) - Q\right)Q^{-1}Y^T\right]_{jj}.
\end{aligned}$$

Since we have $YQ^{-1}Y^T \leq X$, $\max_j X_{jj} \leq \alpha_2$ from (5.15), it follows that if

$$\begin{aligned} &YQ^{-1}(A(k+i)Q + B(k+i)Y - Q)Q^{-1} \\ &\quad \cdot \left(QA^T(k+i) + Y^T B^T(k+i) - Q\right) Q^{-1}Y^T \leq \frac{\beta_2}{\alpha_2} YQ^{-1}Y^T, \quad \text{for all } i \geq 0, \end{aligned} \tag{5.18}$$

then

$$\max_{i \geq 0} \|u(k+i+1|k) - u(k+i|k)\|_\infty^2 \leq \beta_2$$

is satisfied. Furthermore, if

$$\begin{aligned} &\frac{\beta_2}{\alpha_2} Q - (A(k+i)Q + B(k+i)Y - Q)Q^{-1} \\ &\quad \cdot \left(QA^T(k+i) + Y^T B^T(k+i) - Q\right) \geq 0, \qquad \text{for all } i \geq 0, \end{aligned}$$

then (5.18) holds. From Schur complements and $[A(k+i)\ B(k+i)] \in \Omega$, we can see that the above inequality is equivalent to

$$\begin{bmatrix} \beta_2 Q & A_j Q + B_j Y - Q \\ QA_j^T + Y^T B_j^T - Q & \frac{1}{\alpha_2} Q \end{bmatrix} \geq 0, \quad j = 1, 2, \ldots, L.$$

■

5.2.2 Rate Constraints on the Output

In this subsection, we present LMI constraints on rate limits of the output. In the case of l^2 norm, we have the following lemma.

Lemma 5.6 (Rate constraint on the output (l^2 norm case)) *For the system (5.1), if there exist $Q > 0$ and Y satisfying (5.8), (5.9) and*

$$\begin{bmatrix} Q & (A_j Q + B_j Y - Q)^T C \\ C^T(QA_j^T + Y^T B_j^T - Q) & \nu_1 I \end{bmatrix} \geq 0, \quad j = 1, 2, \ldots, L, \tag{5.19}$$

then

$$\max_{i \geq 0} \|y(k+i+1|k) - y(k+i|k)\|_2^2 \leq \nu_1.$$

Proof Following the line of proof of Lemma 5.4. ■

In a similar manner, l^∞ bounds on the output can be translated to sufficient LMI constraints. The development is identical to the preceding development for the l^2 norm constraint if we replace C by C_l and T by T_l, $l = 1, 2, \ldots, n_y$ in (5.19), where

$$y(k) = \begin{bmatrix} y_1(k) \\ \vdots \\ y_{n_y}(k) \end{bmatrix} = Cx(k) = \begin{bmatrix} C_1 \\ \vdots \\ C_{n_y} \end{bmatrix} x(k).$$

Remark 5.2 For the system with structured feedback uncertainty [35], LMI constraints similar to those presented here can be derived by using the $\mathcal{S}$-procedure [60].

5.3 Robust Stability

In this section, we state robust stability of the closed-loop when we use the receding horizon state-feedback control law given in Theorem 5.1 [35].

Theorem 5.2 *Let the state $x(k|k)$ of the uncertain system (5.1) be measurable. Then, the state-feedback matrix F in the control law $u(k+i|k) = Fx(k+i|k), i \geq 0$, which minimizes the upper bound $V(x(k|k))$ on the robust performance objective function at sampling time k and satisfies a set of specified input and output constraints is given by*

$$F = YQ^{-1},$$

where $Q > 0$ and Y are obtained from the solution (if it exists) to the following linear objective minimization problem:

$$\min_{\gamma, Q, Y} \gamma$$

subject to (5.8), (5.9) and the LMIs corresponding to the input and output constraints in lemmas described in the previous sections.

Proof From Lemma 5.1, we know that (5.9) imply that $\mathcal{E}$ is an invariant ellipsoid for the predicted states of the uncertain system (5.1). Hence the arguments in the previous section used to translate the input and output constraints to sufficient LMI constraints hold true. The reset of the proof is similar to the proof of Theorem 5.1. ■

In order to prove robust stability of the closed loop, we need to establish the following lemma [35].

Lemma 5.7 (Feasibility) *Any feasible solution of the optimization in Theorem 5.2 at time k is also feasible for all times $t > k$. Thus, if the optimization problem in Theorem 5.2 is feasible at time k, then it is feasible for all times $t > k$.*

Proof Let us assume that the optimization problem in Theorem 5.2 is feasible at sampling time k. The only LMI in the problem which depends explicitly on the measured state $x(k|k)$ of the system is the following:

$$\begin{bmatrix} 1 & x(k|k)^T \\ x(k|k) & Q \end{bmatrix} \geq 0.$$

Thus, to prove the lemma, we need only prove that this LMI is feasible for all future measured states $x(k+i|k+i), i \geq 1$.

Now, feasibility of the problem at time k implies satisfaction of (5.9), which, using Lemma 5.1, in turn imply that (5.12) is satisfied. Thus, for any $[A(k+i)\ B(k+i)] \in \Omega, i \geq 0$, we must have

$$x(k+i|k)^T Q^{-1} x(k+i|k) < 1, \quad i \geq 1.$$

Since the state measured at $k+1$, that is, $x(k+1|k+1)$, equals $(A(k)+B(k)F)x(k|k)$ for some $[A(k)\ B(k)] \in \Omega$, it must also satisfy this inequality, i.e.,

$$x(k+1|k+1)^T Q^{-1} x(k+1|k+1) < 1,$$

or

$$\begin{bmatrix} 1 & x(k+1|k+1)^T \\ x(k+1|k+1) & Q \end{bmatrix} > 0 \quad \text{(using Schur complements).}$$

Thus the feasibility solution of the optimization problem at time k is also feasible at time $k+1$. Hence the optimization is feasible at time $k+1$. This argument can be continued for time $k+2, k+3, \ldots$ to complete the proof. ∎

Theorem 5.3 (Robust stability) *The feasible receding horizon state-feedback control law obtained from Theorem 5.2 robustly asymptotically stabilizes the closed loop system.*

Proof To prove asymptotic stability, we will establish that $V(x(k|k)) = x(k|k)^T P_k x(k|k)$, where $P_k > 0$ is obtained from the optimal solution at time k, is a strictly decreasing Lyapunov function for the closed-loop.

First, let us assume that the optimization in Theorem 5.2 is feasible at time $k = 0$. Lemma 5.7 then ensures feasibility of the problem at all times $k > 0$. The optimization being convex, therefore, has a unique minimum and a corresponding optimal solution (γ, Q, Y) at each time $k \geq 0$.

Next, we note from Lemma 5.7 that $\gamma, Q > 0, Y$ (or equivalently, γ, $F = YQ^{-1}$, $P = \gamma Q^{-1} > 0$) obtained from the optimal solution at time k are feasible (of course, not necessarily optimal) at time $k+1$. Denoting the values of P obtained from the optimal solutions at time k and $k+1$ respectively by P_k and P_{k+1}, we must have

$$x(k+1|k+1)^T P_{k+1} x(k+1|k+1) \leq x(k+1|k+1)^T P_k x(k+1|k+1). \tag{5.20}$$

This is because P_{k+1} is optimal whereas P_k is only feasible at time $k+1$.

And lastly, we know from Lemma 5.1 that if $u(k+i|k) = F_k x(k+i|k), i \geq 0$ (F_k is obtained from the optimal solution at time k), then for any $[A(k)\ B(k)] \in \Omega$, we must have

$$x(k+1|k)^T P_k x(k+1|k) < x(k|k)^T P_k x(k|k), \qquad (x(k|k) \neq 0) \tag{5.21}$$

(see (5.13) with $i = 0$.)

Since the measured state $x(k+1|k+1)$ equals $(A(k)+B(k)F_k)x(k|k)$ for some $[A(k)\,B(k)] \in \Omega$, it must also satisfy inequality (5.21). Combining this with inequality (5.20) we conclude that

$$x(k+1|k+1)^T P_{k+1} x(k+1|k+1) < x(k|k)^T P_k x(k|k), \qquad (x(k|k) \neq 0).$$

Thus $x(k|k)^T P_k x(k|k)$ is a strictly decreasing Lyapunov function for the closed-loop, which is bounded below by a positive definite function of $x(k|k)$ (see (5.5)). We therefore conclude that $x(k) \to 0$ as $k \to \infty$. ∎

5.4 Numerical Example

In this section, we show an example which illustrates the effectiveness of the presented LMI conditions for rate constraints. Consider the following system [35]

$$\begin{aligned} x(k+1) &= \begin{bmatrix} 1 & 0.1 \\ 0 & 1-0.1\alpha(k) \end{bmatrix} x(k) + \begin{bmatrix} 0 \\ 0.0787 \end{bmatrix} u(k) \\ &:= A(k)x(k) + Bu(k) \\ y(k) &= [1 \quad 0]x(k) := Cx(k) \end{aligned}$$

where $0.1 \leq \alpha(k) \leq 10$. The parameter $\alpha(k)$ is assumed to be arbitrarily time-varying in the indicated range of variation. Then we see that $A(k) \in \Omega = \mathbf{Co}\{A_1, A_2\}$, where

$$A_1 = \begin{bmatrix} 1 & 0.1 \\ 0 & 0.99 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 1 & 0.1 \\ 0 & 0 \end{bmatrix}.$$

Thus the uncertainty set Ω is a polytope, as in (5.2). At each sampling time k, the robust performance objective is as follows:

$$J(k) = \sum_{i=0}^{\infty} \left(y(k+i|k)^2 + 0.00002u(k+i|k)^2 \right).$$

We give an initial state of $x(0) = [1 \ 0]^T$. Here, $\alpha(k)$ is randomly time-varying between 0.1 and 10.

Rate constraint on the input. We first consider the rate constraint on the input. Figure 5.1 shows (a) the amplitudes and (b) the rates of the control inputs of the system. The dashed line shows the input response when only an amplitude constraint $|u(k)|^2 \leq 4$ is imposed. The corresponding LMI is (5.14) with $\alpha_1 = 4$. The solid line shows the input response when the same input amplitude constraint $|u(k)|^2 \leq 4$ and a rate constraint $|u(k+1) - u(k)|^2 \leq 4.5$ are imposed. In this case, the LMI constraints (5.16) with $\beta_1 = 4.5$ as well as (5.14) with $\alpha_1 = 4$ are added to the SDP in Theorem 5.1. Notice that the rate is reduced under the consideration.

Rate constraint on the output Next, we consider the rate constraint on the output. Figure 5.2 shows (a) the amplitudes and (b) the rates of the control outputs of the system. The dashed line shows the output response when no output constraint is imposed. The solid line shows the output response when a rate constraint $|y(k+1) - y(k)|^2 \leq 0.05$ is imposed. The corresponding LMI is (5.19) with $\nu_1 = 0.05$. Notice that the rate is reduced under the consideration.

Rate constraint on the input and output Finally, we consider the rate constraint on both the input and the output. Figure 5.3 shows (a) the rates of the inputs and (b) the rates of the outputs. The dashed lines show the input and output responses when no rate constraint is imposed. The solid lines show the input and output responses when rate constraints $|u(k+1) - u(k)|^2 \leq 4.5$ and $|y(k+1) - y(k)|^2 \leq 0.05$ are imposed. Notice that the rates of both the input and the output are reduced under the consideration.

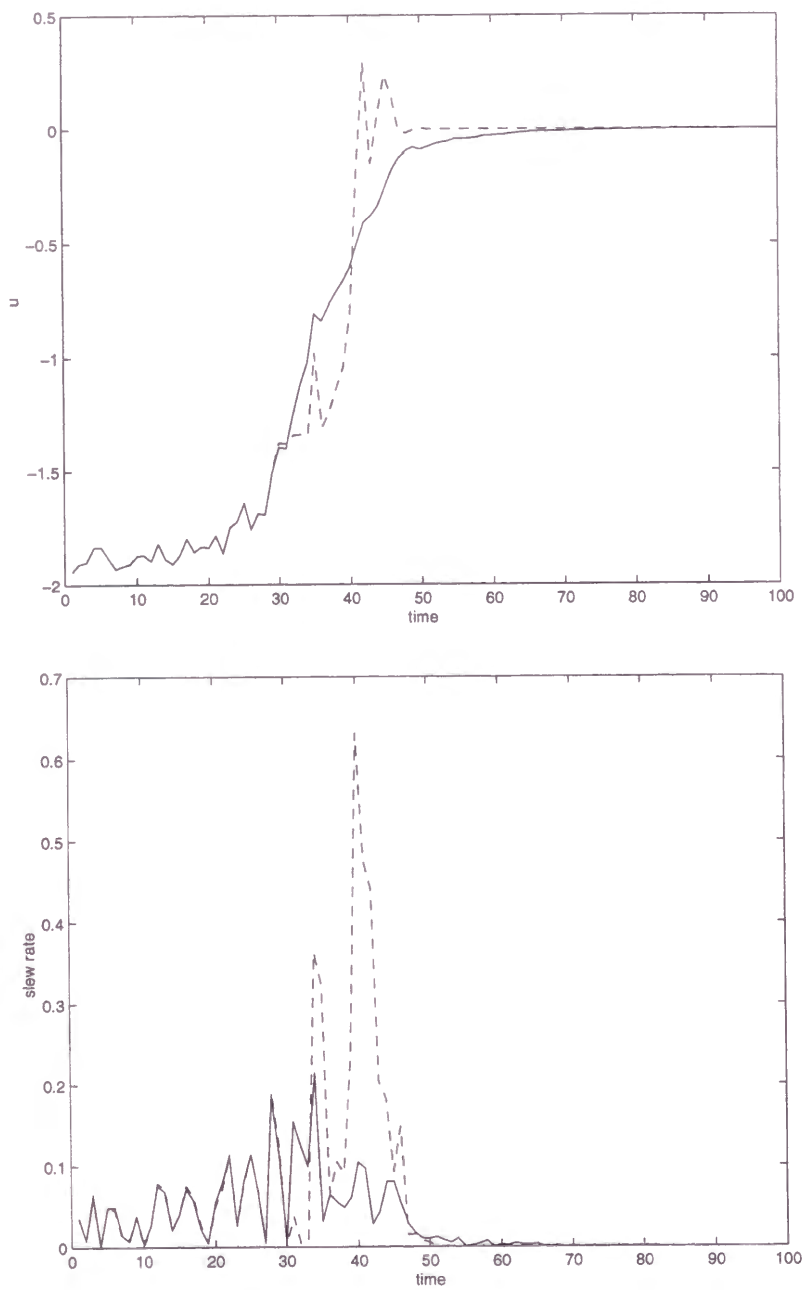

Figure 5.1: Amplitudes and rates of the inputs with rate constraint (solid) and without rate constraint (dashed).

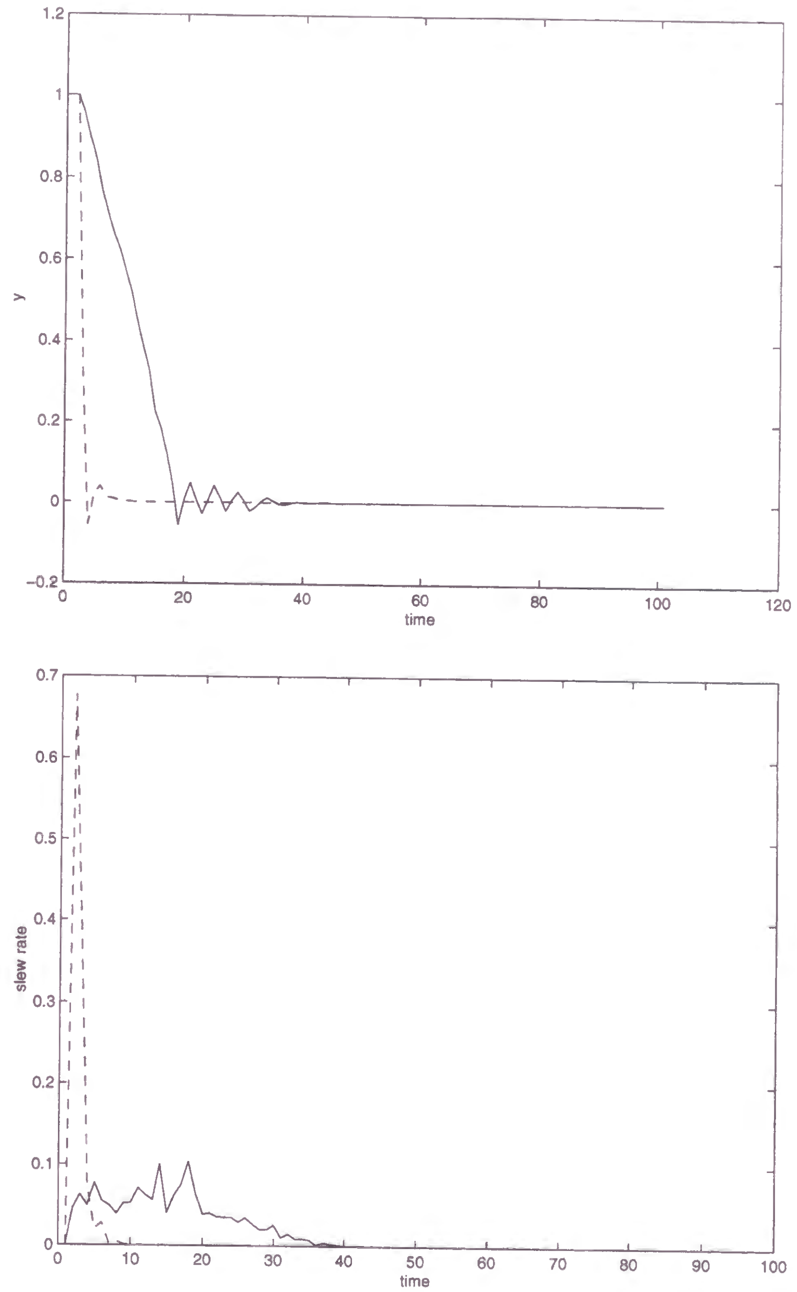

Figure 5.2: Amplitudes and rates of the outputs with rate constraint (solid) and without rate constraint (dashed).

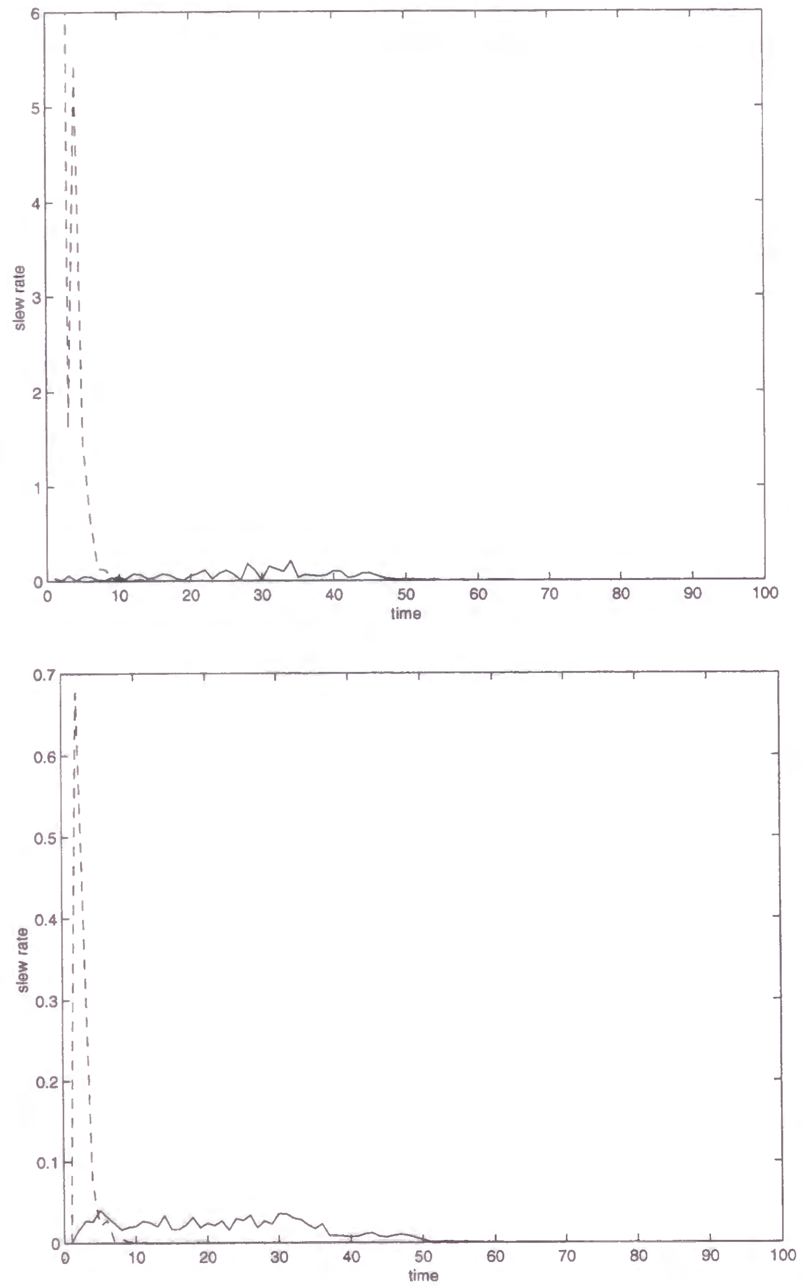

Figure 5.3: Rates of the inputs and the outputs with rate constraints (solid) and without rate constraints (dashed).

Chapter 6

Global Optimization Algorithms for the BMI Problem

Since the so-called LMI approach was studied in the field of control engineering, BMIs have received much attention because BMIs are more general than LMIs and have the ability to naturally characterize many control synthesis problems that are considered to be hard. The following is a list of such problems: the low order controller synthesis [26], multi-objective control and structure [48], distributed control synthesis [29], simultaneous optimization of control and structure [43]. Since Safonov, Goh and others [47, 24] introduced BMIs as a unified description of these various control synthesis problems, some numerical algorithms for BMIs have been investigated intensively with the aim of solving control synthesis problems of practical size [23, 40, 16, 49, 62, 17, 61, 4].

In this chapter, we present a global optimization algorithm for the BMI problem based on the primal-relaxed dual method; this method is a global optimization method for mathematical programming problems whose objective function and constraints are both biconvex [12, 37]. We also modify the algorithm from the viewpoint of computational efficiency. A numerical example is given to illustrate the geometrical interpretation and effectiveness of the proposed method. This chapter is based on the results of [61, 62].

6.1 BMI Problem Formulation

We first formulate the BMI problem. For given matrices $F_{ij} = F_{ij}^T \in \Re^{m\times m}$, $i \in \{0,1,\ldots,n_x\}$, $j \in \{0,1,\ldots,n_y\}$, we define the biaffine function $F : \Re^{n_x} \times \Re^{n_y} \to \Re^{m\times m}$ as follows:

$$F(x,y) := F_{00} + \sum_{i=1}^{n_x} x_i F_{i0} + \sum_{j=1}^{n_y} y_j F_{0j} + \sum_{i=1}^{n_x}\sum_{j=1}^{n_y} x_i y_j F_{ij}. \tag{6.1}$$

We assume that $F(x,y)$ is defined on $X \times Y$, where $X \subset \Re^{n_x}$ and $Y \subset \Re^{n_y}$ denote bounded hyper-rectangles:

$$\begin{aligned} X &:= [x_1^L, x_1^U] \times [x_2^L, x_2^U] \times \cdots \times [x_{n_x}^L, x_{n_x}^U] \\ Y &:= [y_1^L, y_1^U] \times [y_2^L, y_2^U] \times \cdots \times [y_{n_y}^L, y_{n_y}^U] \\ & \quad -\infty < x_i^L < x_i^U < \infty, \quad -\infty < y_j^L < y_j^U < \infty. \end{aligned}$$

Note that suppose

$$F_X(x) := \operatorname{diag}(x_1 - x_1^L, x_1^U - x_1, \ldots, x_{n_x} - x_{n_x}^L, x_{n_x}^U - x_{n_x}),$$

then $X = \{x \mid F_X(x) \geq 0\}$ holds. Here we consider the following optimization problem with a BMI constraint:

Definition 6.1 (BMI Optimization Problem (OP))

$$J_{op} := \min_{x,\lambda,y\in Y} \{\lambda \mid \tilde{F}(x,\lambda,y) \geq 0\},$$

where λ is a scalar variable and $\tilde{F}(x,\lambda,y) := \operatorname{diag}(F(x,y) + \lambda I_m, F_X(x))$.

Remark 6.1 There exists a solution $(x,y) \in X \times Y$ to the BMI $F(x,y) > 0$ if and only if the optimal value of (OP) is negative.

As stated in [23], the primal-relaxed dual approach proposed in [12] cannot be applied to the BMI eigenvalue problem [23, 25, 16, 49]

$$\min_{x\in X, y\in Y} \lambda_{\min}(F(x,y)),$$

since closed form formulae for the gradients of a Lagrangian formulated for this problem are unavailable. However, the BMI optimization problem (OP) defined above has been shown to be solvable by the primal-relaxed dual approach [62, 4].

Basic ideas of primal-relaxed dual approaches date back to the generalized Benders decomposition method [20], which can solve only some special problems. Recently, these approaches have received significant attention in the area of global optimization since Floudas and Visweswaran [12] proposed a global optimization algorithm for a larger class of problems, i.e., mathematical programming problems whose objective and constraints are both biconvex.

In the following three sections, we will present a global optimization method for (OP) based the primal-relaxed dual method [12].

6.2 Primal and Relaxed Dual Problems

In the primal-relaxed dual approach, an original problem is decomposed into a primal problem and a relaxed dual problem which give upper and lower bounds of the optimal value of the original problem, respectively. In this section, we introduce these problems for (OP).

Define the following problem as the primal problem (P):

$$J_p^k := \min_{x,\lambda} \{\lambda \mid \tilde{F}(x,\lambda,y^k) \geq 0\},$$

where k will denote the k-th iteration in the subsequent algorithm and $y^k \in Y$. Since (P) is the problem (OP) where the variable y is fixed to be y^k, it represents an upper bound on the optimal value of (OP), i.e., $J_p^k \geq J_{op}$. Note that (P) is an SDP and therefore can be solved efficiently.

Next, for the purpose of deriving the relaxed dual problem for (OP), we introduce a Lagrangian associated with the problem (OP):

$$L(x,\lambda,y,Z) := \lambda - \mathbf{Tr}\{\tilde{F}(x,\lambda,y)Z\}, \tag{6.2}$$

where $Z = Z^T \in \Re^{(m+2n_x)\times(m+2n_x)}$ is the Lagrange multiplier corresponding to the LMI constraint of the primal problem (P). Then we see that

$$\sup_{Z\geq 0} L(x,\lambda,y,Z) = \begin{cases} \lambda, & \text{if } \tilde{F}(x,\lambda,y) \geq 0 \\ +\infty, & \text{otherwise.} \end{cases}$$

Therefore, for any $y \in Y$, (OP) is rewritten as a min-max problem without the constraint $\tilde{F}(x,\lambda,y) \geq 0$ as follows:

$$\min_{x,\lambda,\tilde{F}(x,\lambda,y)\geq 0} \lambda = \min_{x,\lambda} \max_{Z\geq 0} L(x,\lambda,y,Z)$$

Moreover, since (P) is a convex optimization problem and satisfies Slater's constraint qualification [38], the strong duality theorem [38] holds for (P), i.e., the following relation holds for any $y \in Y$:

$$\min_{x,\lambda} \max_{Z\geq 0} L(x,\lambda,y,Z) = \max_{Z\geq 0} \min_{x,\lambda} L(x,\lambda,y,Z).$$

Hence the problem (OP) is equivalent to the following problem:

$$\min_{y\in Y,\mu} \{\mu \mid \mu \geq \min_{x,\lambda} L(x,\lambda,y,Z) \text{ for all } Z \geq 0\}. \tag{6.3}$$

This problem is very difficult to solve since it contains an infinite number of constraints. To cope with this difficulty, we introduce the following relaxed problem with a finite number of constraints:

$$J_{rd}^K := \min_{y\in Y,\mu} \{\mu \mid \mu \geq \min_{x,\lambda} L(x,\lambda,y,Z^k) \text{ for all } k = 1,\ldots,K\},$$

where $Z^k \geq 0$ is the optimal Lagrange multipliers corresponding to the primal problem (P) for $y = y^k$. This problem is referred to as the relaxed-dual problem (RD) of the original problem (OP). Note that the problem (RD) contains fewer constraints than (6.3), and hence provides a valid lower bound for the original problem (OP), i.e., $J_{rd}^K \leq J_{op}$.

6.3 Properties of the Relaxed Dual Problem

In this section, we discuss properties of the relaxed dual problem (RD) and describe that the properties give a basic idea to solve (RD).

We first give the following property of the Lagrangian $L(x,\lambda,y,Z^k)$:

Property 6.1 *Suppose that Z^k is the optimal Lagrange multipliers corresponding to the problem (P) and that $\mathcal{I}^k$ is the set of i's for which $\nabla_{x_i} L(x,\lambda,y,Z^k)$ is a function of y. Then the following properties hold.*

(i) *Lagrangian $L(x,\lambda,y,Z^k)$ is independent of λ.*

(ii) *If $i \notin \mathcal{I}^k$, i.e., $\nabla_{x_i} L(x,\lambda,y,Z^k)$ is constant, then $\nabla_{x_i} L(x,\lambda,y,Z^k) = 0$ holds.*

Proof

(i) It follows from the Kuhn-Tucker conditions [38] that

$$\nabla_\lambda L(x^k,\lambda^k,y^k,Z^k) = 1 - \mathbf{Tr}\{\mathrm{diag}(I_m, 0_{2n_x})Z^k\} = 0.$$

Therefore

$$\begin{aligned} L(x,\lambda,y,Z^k) &= (1 - \mathbf{Tr}\{\mathrm{diag}(I_m, 0_{2n_x})Z^k\})\lambda - \mathbf{Tr}\{\mathrm{diag}(F(x,y), F_X(x))Z^k\} \\ &= -\mathbf{Tr}\{\mathrm{diag}(F(x,y), F_X(x))Z^k\}. \end{aligned}$$

(ii) The property is obvious since $\nabla_{x_i} L(x,\lambda,y^k,Z^k) = 0$ holds from the Kuhn-Tucker conditions.

■

Next, we will focus our attention on the problem (RD). The problem (RD) is still difficult to solve since it contains an inner minimization problem which is denoted as an inner relaxed dual problem (IRD):

$$\min_{x,\lambda} L(x,\lambda,y,Z^k).$$

However, using special properties such as Property 6.1, we can solve the problem (RD) by decomposing it into tractable subproblems which contains no inner minimization problems [62, 4]. The following property of (IRD) enables such a decomposition [62].

Property 6.2 *Let $B_j, j = 1,\ldots,2^{n_x}$ indicate combinations of lower/upper bounds of the variables $x_i, i = 1,\ldots,n_x$, e.g., $x^{B_1} = (x_1^L, x_2^L, x_3^U, \ldots, x_{n_x}^L)^T$ for $B_1 = (L,L,U,\ldots,L)$, and let $\mathcal{B} := \{B_1,\ldots,B_{2^{n_x}}\}$ be the set of all bound combinations. Suppose that $\bar{x}^k(y)$ is a function of y such that, for every $y \in Y$,*

$$L(\bar{x}^k(y),\lambda,y,Z^k) := \min_x L(x,\lambda,y,Z^k).$$

Then, for every $y \in Y$,

$$
\begin{aligned}
&L(\bar{x}^k(y), \lambda, y, Z^k) \\
&= \left\{ \begin{array}{l}
\min\limits_{B_j \in \mathcal{B}} L(x^{B_j}, \lambda, y, Z^k) \\
\text{subject to} \ \left. \begin{array}{l} \nabla_{x_i} L(x, \lambda, y, Z^k) \geq 0, \text{ if } x_i^{B_j} = x_i^L \\ \nabla_{x_i} L(x, \lambda, y, Z^k) \leq 0, \text{ if } x_i^{B_j} = x_i^U \end{array} \right\} \text{for all } i \in \mathcal{I}^k
\end{array} \right\}.
\end{aligned}
\tag{6.4}
$$

Proof In Property 4.2 in [12], a biconvex Lagrangian is considered, so that an inequality sign "$\geq$" holds instead of the equality sign in (6.4). In our case, however, Lagrangian (6.2) is bilinear over $X \times Y$. Noting this bilinearity, we can complete the proof in a similar way. ■

Suppose that the relaxed-dual problem (RD) has one inequality constraint and that the optimal solution is (x^*, y^*). Then, Property 6.2 states that the optimal solution x^* is achieved at an extreme point of X and the optimal solution y^* is obtained by solving linear programs for all extreme points of X. We will illustrate this fact in the subsequent numerical example.

Remark 6.2 From Property 6.2, we see that we can solve the problem (RD) with one more constraint by solving some linear programs. Note that, however, the number of the linear programs is often huge, and therefore, solving all linear programs is not efficient from the computational point of view. In [12], to cope with the difficulty, the lower bound of the optimal value of (RD) given by fewer linear programs is used in the global optimization algorithm of (OP).

The following property shows that a lower bound of the optimal value of the problem (RD) is obtained from some subproblems.

Property 6.3 *At the K-th iteration,*

(i) define $\mathcal{J}(k, K), k < K$ to be the set of j's such that

$$
\left. \begin{array}{l}
\nabla_{x_i} L(x, \lambda, y^K, Z^k) \geq 0, \text{if } x_i^{B_j} = x_i^L \\
\nabla_{x_i} L(x, \lambda, y^K, Z^k) \leq 0, \text{if } x_i^{B_j} = x_i^U
\end{array} \right\} \text{for all } i \in \mathcal{I}^k;
$$

(ii) define $\mu^{\text{stor}}(K, B_l)$ and $y^{\text{stor}}(K, B_l)$ to be the optimal solutions of the following subproblem (SUBP1) associated with (RD):

$$
\begin{aligned}
&\mu^{\text{stor}}(K, B_l) \\
&= \left\{ \begin{array}{l}
\min\limits_{y \in Y, \mu} \mu \\
\text{subject to} \\
\left. \begin{array}{l}
\mu \geq L(x^{B_j}, \lambda, y, Z^k) \\
\quad \left. \begin{array}{l} \nabla_{x_i} L(x, \lambda, y, Z^k) \geq 0, \text{if } x_i^{B_j} = x_i^L \\ \nabla_{x_i} L(x, \lambda, y, Z^k) \leq 0, \text{if } x_i^{B_j} = x_i^U \end{array} \right\} \text{for all } i \in \mathcal{I}^k
\end{array} \right\} \begin{array}{l} \text{for } j = \mathcal{J}(k, K), \\ k = 1, \dots, K-1 \end{array} \\
\text{and } \mu \geq L(x^{B_l}, \lambda, y, Z^K) \\
\quad \left. \begin{array}{l} \nabla_{x_i} L(x, \lambda, y, Z^K) \geq 0, \text{if } x_i^{B_l} = x_i^L \\ \nabla_{x_i} L(x, \lambda, y, Z^K) \leq 0, \text{if } x_i^{B_l} = x_i^U \end{array} \right\} \text{for all } i \in \mathcal{I}^K.
\end{array} \right.
\end{aligned}
$$

Then, for the optimal value μ_{RD}^K of the problem (RD) and $\mu_{\min}^K := \min_{B_l \in \mathcal{B}} \mu^{\text{stor}}(K, B_l)$, $\mu_{\min}^K \leq \mu_{\text{RD}}^K$ holds.

Proof This property corresponds to Property 4.4 in [12]. According to the context in [12], it follows the result since Property 6.2 is more special than Property 4.2 in [12]. ■

Remark 6.3 Note that $\mathcal{J}(k, K)$ is uniquely determined for any combination of K and k $(k < K)$ for the BMI problem (OP). Also note that (SUBP1) is a linear program with variables y and μ.

6.4 Primal-Relaxed Dual Algorithm for BMIs

In this section, we present a global optimization algorithm for the BMI problem based on the primal-relaxed dual method. Roughly speaking, the algorithm consists of the following two procedures:

(i) Solve the primal problem (P) for $y = y^k$ and update the obtained upper bound.

(ii) Solve the subproblems of the relaxed dual problem (RD) and update the obtained lower bound.

These two procedures are to be repeated until the difference between the upper and lower bounds becomes less than the prescribed tolerance $\epsilon(> 0)$.

Algorithm 6.1 (BMI Primal-Relaxed Dual Algorithm)

Step 0. Initialization of Parameters.

Let P^{UBD} and M^{LBD} be a very large *positive* number and a very large *negative* number, respectively. Select a convergence tolerance parameter ϵ (> 0). Set $K = 1$ and select an initial fixed value $y^1 \in Y$.

Step 1. Primal Problem.

Store the value of y^K. Solve the primal problem (P) for $y = y^K$. Store the optimal Lagrange multiplier Z^K. Update the upper bound so that

$$P^{\text{UBD}} = \min\,(P^{\text{UBD}}, P^K)$$

where P^K is the solution of the K-th primal problem.

Step 2. Selection of Lagrangians from the Previous Iterations.

For $k = 1, 2, \ldots, K-1$, select the Lagrangian corresponding to $j = \mathcal{J}(k, K)$.

Step 3. Relaxed Dual Problem.

For all $B_l \in \mathcal{B}$, solve the subproblem (SUBP1) and store the solutions $\mu^{\text{stor}}(K, B_l)$, $y^{\text{stor}}(K, B_l)$.

Step 4. Selection of a New Lower Bound and y^{K+1}.

From the stored set μ^{stor}, select the minimum μ^K_{min}, and set $M^{\text{LBD}} = \mu^K_{\text{min}}$. Also, select the corresponding stored value of y^{stor} as y^{K+1}. Delete μ^K_{min} and y^{K+1} from μ^{stor} and y^{stor}, respectively.

Step 5. Check for Convergence.

Check if $P^{\text{UBD}} - M^{\text{LBD}} \leq \epsilon$. If yes, stop, else set $K = K + 1$, and return to Step 1.

Remark 6.4 In Step 3 of the algorithm, the relaxed-dual problem (RD) is decomposed into the subproblems (SUBP1) that are formulated as linear programming problems. By solving these subproblems, we obtain a lower bound on the optimal value μ^K_{RD} of the original problem (RD), i.e., μ^K_{min} computed in Step 4 is a lower bound of μ^K_{RD}. Note that the convergence and global optimality of the algorithm is proved in [12, 36] although it should be noted that (RD) is not solved exactly.

Remark 6.5 Note that the optimal Lagrange multiplier Z^k (i.e., the dual optimal solution) as well as the optimal solution of the problem (P) can be computed by using computer programs developed to solve SDPs based on the primal-dual interior-point method such as [55]. Also note that since (P) is always feasible for any $y^k \in Y$, we need not cope with the case where (P) is infeasible as in [12].

6.5 Modification of the Algorithm

In this section, we present a modified algorithm from the viewpoint of practical efficiency.

We see that the number of constraints of the subproblem (SUBP1) in Step 3 is $O(K)$, i.e., it grows linearly with the number of iterations K. In many numerical experiments, this increase in the number of the constraints leads to inefficiency in solving one subproblem, so that the algorithm requires a very long time to terminate. Also, some of these constraints are often redundant as shown in the subsequent numerical example. Therefore, one method for improving computational efficiency is to select only active constraints of the previous subproblem in Step 2. In such a selection of Lagrangians, one subproblem contains a constant number of constraints, and therefore, we can expect that whole efficiency of the algorithm is improved. In view of this observation, we propose a new method for selecting Lagrangians in accordance with the following definition:

Definition 6.2 At the Kth iteration, define $\mathcal{J}_{\text{act}}(K)$ to be the set of (j, k)'s, $k < K$ such that

$$\mu^{K-1}_{\text{min}} = L(x^{B_j}, \lambda, y^K, Z^k), \tag{6.5}$$

i.e., $\mathcal{J}_{\text{act}}(K)$ is the set of (j, k)'s corresponding to the active constraints of the subproblems at the previous iteration.

By the above selection method of Lagrangians, Algorithm 6.1 is modified in Step 2 and Step 3 as follows:

Algorithm 6.2 (Modified BMI Primal-Relaxed Dual Algorithm)

Step 0. Initialization of Parameters.

Let P^{UBD} and M^{LBD} be a very large *positive* number and a very large *negative* number, respectively. Select a convergence tolerance parameter ϵ (> 0). Set $K = 1$ and select an initial fixed value $y^1 \in Y$.

Step 1. Primal Problem.

Store the value of y^K. Solve the primal problem (P) for $y = y^K$. Store the optimal Lagrange multiplier Z^K. Update the upper bound so that

$$P^{\text{UBD}} = \min\left(P^{\text{UBD}}, P^K\right)$$

where P^K is the solution of the K-th primal problem.

Step 2. Selection of Lagrangians from the Previous Iterations.

Select the Lagrangians corresponding to $\mathcal{J}_{\text{act}}(K)$ determined at the $K-1$th iteration.

Step 3. Relaxed Dual Problem.

For all $B_l \in \mathcal{B}$, solve the subproblem (SUBP2)

$$\mu^{\text{stor}}(K, B_l) = \left\{ \begin{array}{l} \min\limits_{y \in Y, \mu} \mu \\ \text{subject to} \\ \quad \left. \begin{array}{l} \mu \geq L(x^{B_j}, \lambda, y, Z^k) \\ \left. \begin{array}{l} \nabla_{x_i} L(x, \lambda, y, Z^k) \geq 0, \text{if } x_i^{B_j} = x_i^L \\ \nabla_{x_i} L(x, \lambda, y, Z^k) \leq 0, \text{if } x_i^{B_j} = x_i^U \end{array} \right\} \text{for all } i \in \mathcal{I}^k \end{array} \right\} \text{for all } (j,k) \in \mathcal{J}_{\text{act}}(K) \\ \quad \text{and } \mu \geq L(x^{B_l}, \lambda, y, Z^K) \\ \quad \left. \begin{array}{l} \nabla_{x_i} L(x, \lambda, y, Z^K) \geq 0, \text{if } x_i^{B_l} = x_i^L \\ \nabla_{x_i} L(x, \lambda, y, Z^K) \leq 0, \text{if } x_i^{B_l} = x_i^U \end{array} \right\} \text{for all } i \in \mathcal{I}^K \end{array} \right.$$

and store the solutions $\mu^{\text{stor}}(K, B_l)$, $y^{\text{stor}}(K, B_l)$.

Step 4. Selection of a New Lower Bound and y^{K+1}.

From the stored set μ^{stor}, select the minimum μ^K_{min}, and set $M^{\text{LBD}} = \mu^K_{\text{min}}$. Also, select the corresponding stored value of y^{stor} as y^{K+1}. Delete μ^K_{min} and y^{K+1} from μ^{stor} and y^{stor}, respectively.

Step 5. Check for Convergence.

Check if $P^{\text{UBD}} - M^{\text{LBD}} \leq \epsilon$. If yes, stop, else set $K = K + 1$, and return to Step 1.

Remark 6.6 Only active constraints of the subproblem (SUBP2) at the previous iteration $K - 1$ are selected as constraints of the subproblem (SUBP2) at the iteration K. Since the number of the active constraints is independent of K, the number of constraints selected from $\mathcal{J}_{\text{act}}$ is independent of K. For this modified algorithm, the convergence and global optimality are proved as in [12, 37].

Remark 6.7 The lower bound computed by Algorithm 6.2 is less than or equal to the lower bound computed by Algorithm 6.1 at every iteration, i.e., Algorithm 6.2 is more relaxed than Algorithm 6.1. On the other hand, Algorithm 6.2 requires less computation time than Algorithm 6.1 at every iteration. While we cannot conclude which of the algorithms is better from the viewpoint of theoretically computational efficiency, we can expect that practical efficiency for many BMI problems is improved as illustrated in the subsequent numerical example.

6.6 Branch and Bound Algorithm for BMIs

In this section, we present a global optimization algorithm for BMIs based on the branch and bound approach [23]. So far, some improvements for BMI branch and bound algorithms have been discussed [16, 49, 18]. We here show one of the improved branch and bound algorithms.

In the branch and bound approach, the following problem is considered in order to solve the BMI $F(x, y) > 0$.

Definition 6.3 (BMI Eigenvalue Problem) Given the hyper-rectangle $Q := X \times Y$ and the function $F : Q \to \Re^{m\times m}$ defined by (6.1), define

$$\Lambda(x,y) := -\lambda_{\min}(F(x,y)). \tag{6.6}$$

The BMI eigenvalue minimization problem is

$$\min_{(x,y)\in Q} \Lambda(x,y)$$

Let $\Phi(Q)$ denote the optimal value function of the BMI eigenvalue minimization problem, i.e., $\Phi(Q) := \min_{(x,y)\in Q} \Lambda(x,y)$. Clearly, there exists a solution $(x,y) \in Q$ to BMI (6.1) if and only if $\Phi(Q) < 0$. The basic requirement for a branch and bound algorithm is the existence of two functions, Φ_L and Φ_U, on the family of hyper-rectangles Q such that the following conditions hold [2, 3]:

C1. $\Phi_L(Q)$ gives a lower bound and $\Phi_U(Q)$ gives an upper bound on $\Phi(Q)$, i.e.,

$$\Phi_L(Q) \le \Phi(Q) \le \Phi_U(Q)$$

for every hyper-rectangle Q.

C2. Let Size(Q) denote the length of the longest side of the hyper-rectangle Q, then as Size$(Q) \to 0$, $\Phi_U(Q) - \Phi_L(Q) \to 0$ uniformly.

It is shown that the branch and bound algorithm converges in a finite number of steps if the the conditions C1 and C2 are satisfied [3]. Several types of $\Phi_L(Q)$ and $\Phi_U(Q)$ satisfying these conditions have been given in [23, 16, 49, 18]. We present simple functions among them as below.

A simple $\Phi_L(Q)$ [23] is given by

$$\Phi_L(Q) := \min_{\lambda,(x,y)\in Q, W\in\mathcal{W}} \{\lambda \mid \hat{F}_L(x,\lambda,y,W) \ge 0\}, \tag{6.7}$$

where

$$\hat{F}_L(x,\lambda,y,W) := F_L(x,y,W) + \lambda I,$$
$$F_L(x,y,W) := F_{00} + \sum_{i=1}^{n_x} x_i F_{i0} + \sum_{j=1}^{n_y} y_j F_{0j} + \sum_{i=1}^{n_x}\sum_{j=1}^{n_y} w_{ij} F_{ij},$$

$$\mathcal{W} := \{W \in \Re^{n_x \times n_y} \mid w_{ij}^L \le w_{ij} \le w_{ij}^U\},$$
$$w_{ij}^L := \min\{x_i^L y_j^L, x_i^L y_j^U, x_i^U y_j^L, x_i^U y_j^U\},$$
$$w_{ij}^U := \max\{x_i^L y_j^L, x_i^L y_j^U, x_i^U y_j^L, x_i^U y_j^U\}.$$

A simple $\Phi_U(Q)$ [16] is given by

$$\Phi_U(Q) := -\lambda_{\min}(F(x^*, y^*)),$$

where (x^*, y^*) is the optimal solution to the minimization problem (6.7).

Note that $\Phi_L(Q)$ is computed by solving the SDP whose feasible region is enlarged in comparison with the equivalent SDP to the problem (6.6), and therefore, $\Phi_L(Q)$ gives a lower bound of $\Phi(Q)$.

Once we have the functions $\Phi_U(Q)$ and $\Phi_L(Q)$ as shown above, it is straightforward to adapt the branch and bound algorithm given in [2, 3] to globally minimize $\Phi(Q)$.

Algorithm 6.3 (BMI Branch and Bound Algorithm)

Step 0. Fix $\epsilon > 0$. Set $k := 0$, $Q_0 := Q$, $S_0 := Q_0$. $L_0 := \Phi_L(Q_0)$, $U_0 := \Phi_U(Q_0)$.

Step 1. Select $\bar{Q}$ from S_k such that $L_k = \Phi_L(\bar{Q})$. $S_{k+1} := S_k - \{\bar{Q}\}$.

Step 2. Split $\bar{Q}$ along its longest edge into $\bar{Q}_1$ and $\bar{Q}_2$.

Step 3. For $i = 1, 2$, if $\Phi_L(\bar{Q}_i) \le U_k$, compute $\Phi_U(\bar{Q}_i)$ and set $S_{k+1} := S_{k+1} \cup \{\bar{Q}_i\}$.

Step 4. $U_{k+1} := \min_{Q\in S_{k+1}} \Phi_U(Q)$.

Step 5. Pruning: $S_{k+1} := S_{k+1} - \{Q \mid \Phi_L(Q) > U_{k+1}\}$.

Step 6. $L_{k+1} := \min_{Q\in S_{k+1}} \Phi_L(Q)$.

Step 7. Check if $U_k - L_k < \epsilon$. If yes, stop, else set $k := k + 1$, and return to Step 1.

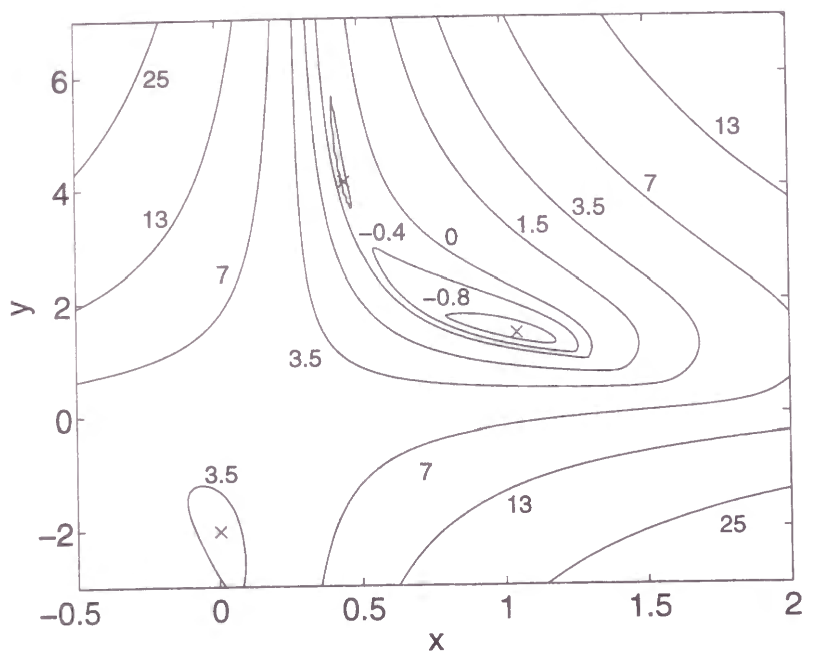

Figure 6.1: Contour plot of $-\lambda_{\min}(F(x,y))$

6.7 Numerical Example

Consider the BMI optimization problem with the following BMI [23, 16, 33]:

$$\begin{aligned} F(x,y) &= \begin{bmatrix} 10 & 0.5 & 2 \\ 0.5 & -4.5 & 0 \\ 2 & 0 & 0 \end{bmatrix} + x \begin{bmatrix} -9 & -0.5 & 0 \\ -0.5 & 0 & 3 \\ 0 & 3 & 1 \end{bmatrix} \\ &\quad + y \begin{bmatrix} 1.8 & 0.1 & 0.4 \\ 0.1 & -1.2 & 1 \\ 0.4 & 1 & 0 \end{bmatrix} + xy \begin{bmatrix} 0 & 0 & -2 \\ 0 & 5.5 & -3 \\ -2 & -3 & 0 \end{bmatrix} > 0, \\ X \times Y &= [-0.5, 2] \times [-3, 7]. \end{aligned}$$

Figure 6.1 shows the contour plots of $-\lambda_{\min}(F(x,y))$. It should first be noted that there are three local minima shown by $\times$, and that $(x,y) = (1.05, 1.42)$ gives the global minimum. Obviously, this optimization problem is non-convex.

We solve this problem by the three types of global optimization algorithms:

- PRD1: the primal-relaxed dual method (Algorithm 6.1);

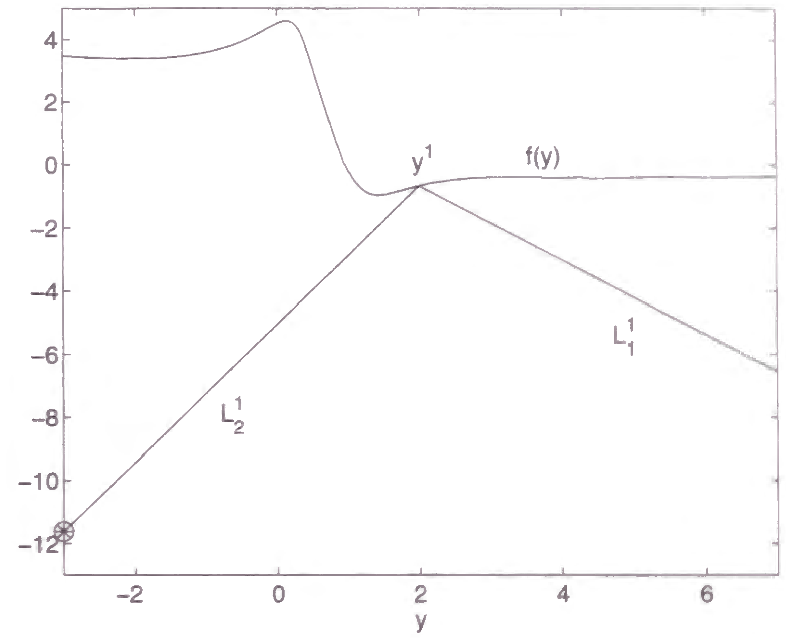

Figure 6.2: Lagrangians and $f(y)$ at the iteration $K = 1$

- PRD2: the modified primal-relaxed dual method (Algorithm 6.2);
- BB: the branch and bound method (Algorithm 6.3).

Before we mention the numerical results, we give the geometrical interpretation of the primal-relaxed dual algorithms PRD1 and PRD2 by using the above example. The Lagrangians and

$$f(y) := -\min_{x,\lambda}\{\lambda \mid \tilde{F}(x,\lambda,y) \geq 0\}$$

at iterations $K = 1$–4 are shown in Figures 6.2–6.5, respectively, where L_j^K is the jth Lagrangian at the Kth iteration, i.e, $L_j^K := L(x^{B_j}, \lambda, y, Z^K)$. In Figures 6.2–6.5, $*$ and $\circ$ denote the pairs $(y^{K+1}, \mu_{\min}^K)$ computed by the algorithms PRD1 and PRD2, respectively.

Now we give a geometrical explanation of the algorithm from the first to the fourth iteration.

Iteration 1. Consider the starting point of $y^1 = 2$. By solving the primal problem (P) for $y^1 = 2$ at $K = 1$, we have the optimal solution and the Lagrange multiplier Z^1. Note that the subproblems solved in Step 3 are same in both algorithms PRD1 and PRD2.

The subproblem for $x^{B_1} = -0.5$ is

$$\min_{y \in Y, \mu} \{\mu \mid \mu \geq L_1^1(y), y - y^1 \geq 0\},$$

and the subproblem for $x^{B_2} = 2$ is

$$\min_{y \in Y, \mu} \{\mu \mid \mu \geq L_2^1(y), y - y^1 \leq 0\}.$$

The optimal solutions of these two subproblems are $(7, -6.5245)$ and $(-3, -11.5853)$ for (y, μ). Thus, after the first iteration, there are two solutions of (μ, y) in the stored set. From these solutions, the solution corresponding to the minimum μ is chosen. In this case, the solution $(y, \mu^1_{\text{min}}) = (-3, -11.5853)$ is chosen. Hence the fixed value for the second iteration is $y^2 = -3$. The selected solution is then deleted from the stored set. Note that the piecewise linear function formed from the Lagrangians L_1^1, L_2^1 in Figure 6.2 is (6.4) in Property 6.2, i.e.,

$$L(\bar{x}^1(y), \lambda, y, Z^1) = \left\{ \begin{array}{l} \min\limits_{B_j \in \{L,U\}} L(x^{B_j}, \lambda, y, Z^1) \\ \text{subject to} \quad \nabla_x L(x, \lambda, y, Z^1) \geq 0, \text{ if } x^{B_j} = x^L, \\ \qquad\qquad\quad \nabla_x L(x, \lambda, y, Z^k) \leq 0, \text{ if } x^{B_j} = x^U \end{array} \right\}, \tag{6.8}$$

and therefore, $L(\bar{x}(y), \lambda, y, Z^1)$ is minimized in Step 3, 4 in the algorithm PRD1 and PRD2 by solving the above two subproblems. At the first iteration, the results of these algorithms are completely same.

Iteration 2. Solve the primal problem (P) for $y^2 = -3$. Then we have the Lagrange multiplier Z^2 and update the upper bound. The algorithm PRD1 gives $\mathcal{J}(1,2) = 2$ and the algorithm PRD2 gives $\mathcal{J}_{\text{act}}(2) = (2,1)$. Thus the Lagrangians selected in Step 2 are same in both algorithms, and therefore, the subproblems of (RD) solved in both algorithms are same at the second iteration. The subproblem for $x^{B_1} = -0.5$ is

$$\min_{y \in Y, \mu} \{\mu \mid \mu \geq L_2^1(y), y - y^1 \leq 0, \mu \geq L_1^2(y), y - y^2 \leq 0\},$$

and the solution is $(y, \mu) = (-3, 3.4757)$. The subproblem for $x^{B_2} = 2$ is

$$\min_{y \in Y, \mu} \{\mu \mid \mu \geq L_2^1(y), y - y^1 \leq 0, \mu \geq L_2^2(y), y - y^2 \geq 0\},$$

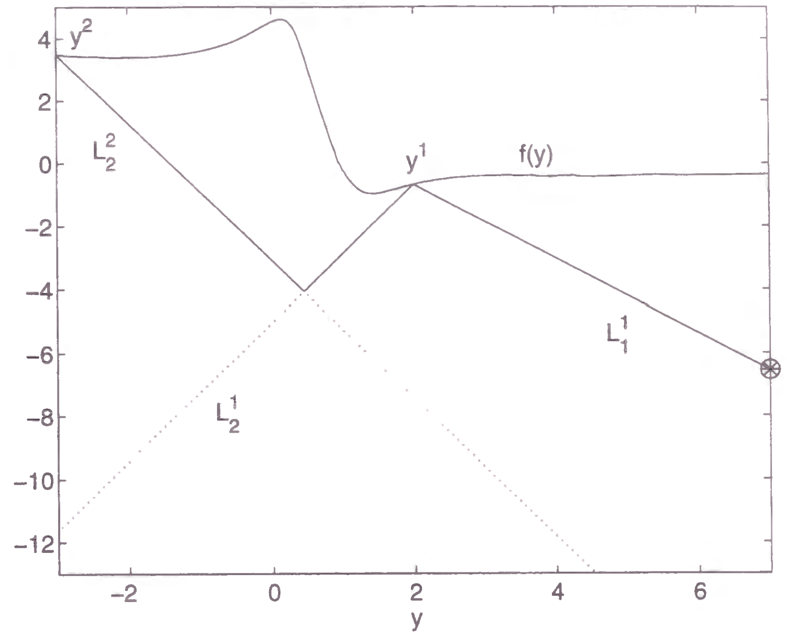

Figure 6.3: Lagrangians and $f(y)$ at the iteration $K = 2$

and the solution is $(y, \mu) = (0.4450, -4.0573)$. From the set which stores these solutions, the lower bound in Step 4 is updated as $\mu^2_{\text{min}} = -6.5245$, and $y^3 = 7$ is selected for the next iteration. Note that minimization in Steps 3 and 4 is performed over the piecewise linear function shown by the solid line in Figure 6.3.

Iteration 3. The algorithm PRD1 gives $\mathcal{J}(1,3) = 1$ and $\mathcal{J}(2,3) = 2$, and the algorithm PRD2 gives $\mathcal{J}_{\text{act}}(3) = (1,1)$. Thus the number of constraints of the subproblems in the algorithm PRD2 is less than in the algorithm PRD1 by the constraints related to L_2^2. Note that, however, the results solved in both algorithms give the same $\mu^3_{\text{min}} = -4.0620$. Therefore we see that the subproblems in Algorithm 6.1 includes some redundant constraints. For reference, we show the optimal solution of the problem (RD) as $\diamond$ in Figure 6.4. Notice that the lower bound of the optimal value of (RD) is obtained at the third iteration while (RD) is exactly solved at the first and second iterations.

Iteration 4. At the forth iteration, the piecewise linear function formed from the Lagrangians in the algorithm PRD1 is different from that in the algorithm PRD2. These piecewise linear functions are shown in Figure 6.5. The lower bound in the algorithm PRD1 is larger than that in the algorithm PRD2 because the constraint of L_2^3 is active in

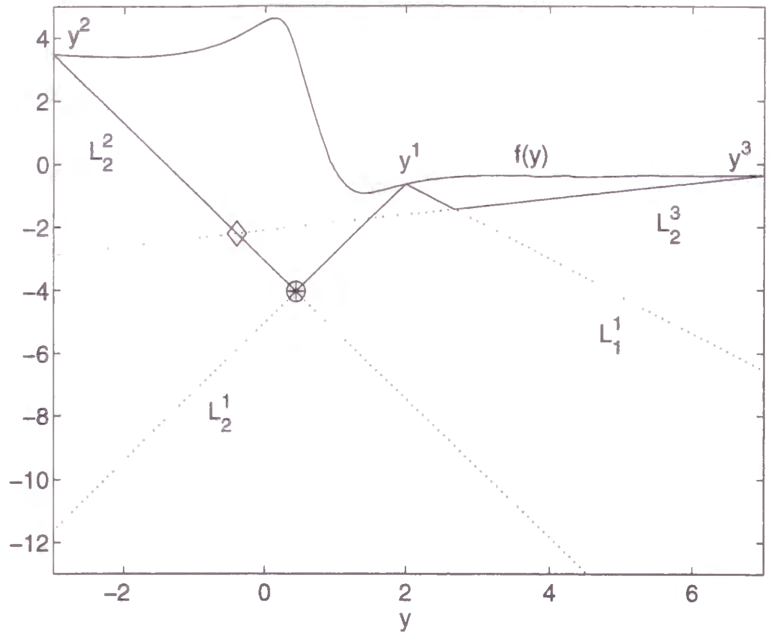

Figure 6.4: Lagrangians and $f(y)$ at the iteration $K = 3$

the algorithm PRD1.

From the above explanation at each iteration, we see that the function $f(y)$ is underestimated by the Lagrangians.

In the following, we present the numerical results. For the tolerance $\epsilon = 10^{-5}$, performance of three global optimization algorithms is shown in Table 6.1. For all computation, MATLAB software is used on Sun SPARCstation 20. Also, the primal problem (P) is solved by SP (MATLAB Version) [55], and the command `lp` in the MATLAB optimization toolbox is performed for the subproblems of (RD). It is seen from Table 6.1 that the algorithms PRD1 and PRD2 is more efficient than the algorithm BB. Also note that

Table 6.1: Performance of algorithms

	PRD1	PRD2	BB
CPU time (sec.)	4.6	2.8	105.5
Number of iterations	29	26	815

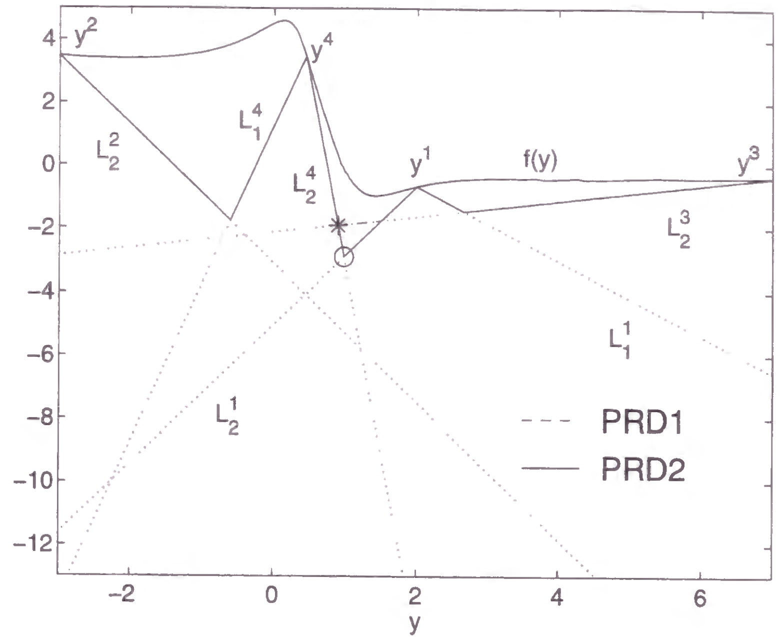

Figure 6.5: Lagrangians and $f(y)$ at the iteration $K = 4$

Table 6.2: Performance of algorithms

	PRD1	PRD2	BB
Average CPU time (sec.)	6.3	3.3	340.7
Average number of iterations	29.2	29.3	1512.8

the algorithms PRD2 terminate faster than the algorithm PRD1, while the number of iterations in the algorithm PRD2 is slightly more than that in the algorithm PRD1. The convergence of lower/upper bounds in the algorithms PRD1 and PRD2 are shown in Figure 6.6.

To illustrate the effectiveness of the proposed method for other examples, we show the average results for the problem with 10 BMIs whose coefficient matrices of $F(x, y)$ are randomly generated. Here all conditions are as same as those in the above example. Table 6.2 shows the average performance for these problems.

Note that we have the almost same result as Table 6.1, i.e., the proposed algorithms

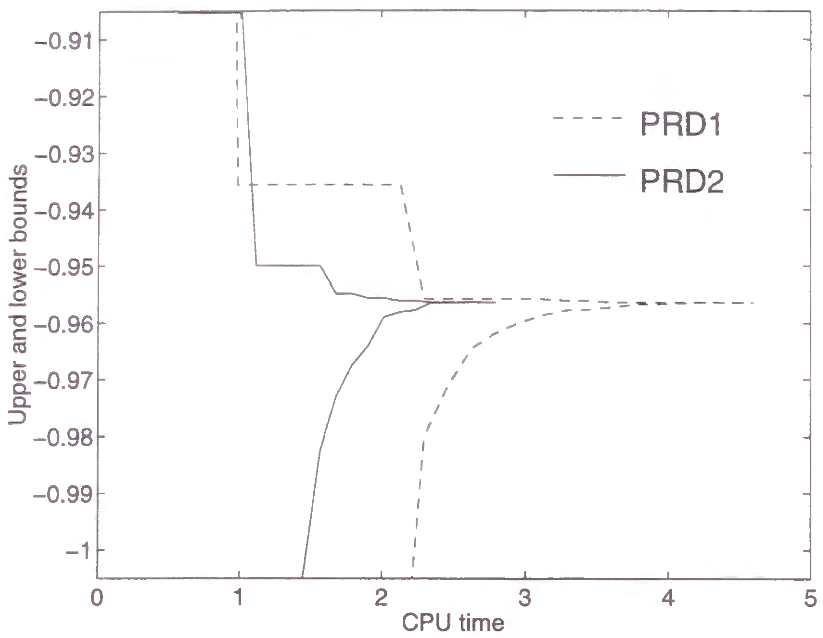

Figure 6.6: Upper and lower bounds

PRD1 and PRD2 are more efficient than the algorithm BB, and the algorithm PRD2 terminates faster than the algorithm PRD1.

Chapter 7

Conclusion

In this thesis, we have investigated control system analysis and synthesis based on LMIs and an optimization method for solving BMIs. We state the conclusions of each chapters in the following.

In **Chapter 3**, we have presented a technique for estimating the sensitivity of the optimal solution and the optimal value function with respect to perturbations in SDP. Also, we have applied this result to control systems with parametric uncertainties. Although the numerical example was restricted to the case of the H^∞ norm, our result can be applied to the problems including the other design objectives, e.g., the H^2 norm and the γ-entropy, which are formulated as SDP.

In **Chapter 4**, we have shown that control system design problem with the tradeoff between evaluated uncertainty ranges and control performance is reducible to an optimization problem with an infinite number of BMI constraints. An approximate method for the problem has been presented, and its convergence has been proved. In a numerical example, this method has given a better performance than standard robust control methods from the viewpoints of the tradeoff.

Chapter 5 has presented the LMI conditions for the rate constraints of the input and output in the framework of the robust LMI-based MPC [35]. A numerical example has shown that the MPC method with these input and output constraints provides a good performance, although the presented LMIs are sufficient conditions for the rate limits. It should be noted that the presented LMI conditions can be also used in the framework of the standard state-feedback synthesis.

In **Chapter 6**, we have presented a global optimization algorithm for the BMI problem based on the primal-relaxed dual method, and modified this algorithm from the viewpoint of computational efficiency. We also have described the branch and bound algorithm for BMIs, and have compared it with the presented methods. Numerical examples have been given to illustrate the geometrical interpretation and effectiveness of the proposed method.

We now discuss future directions of our research. The numerical examples in Chapter 6 have shown that only very small BMIs can be solved in practical time. In fact, however, further improvement or alternative directions are required to solve BMIs corresponding to control problems of practical size. From this point of view, two directions should be considered as follows.

The first direction is to develop a method for solving a more restrictive class of BMIs that arises in control. In fact, most of BMIs that arise in control are more restrictive than the BMI formulated in this thesis. As a study related to this direction, MPEP (Matrix Product Eigenvalue Problems) was proposed [65]. The MPEP is a subclass of BMI problem and can express even control problems which cannot (or are not likely to) be cast into LMIs. At present, a method for solving the MPEP is proposed but it is not still practical for control problems of practical size [65].

The second direction is a probabilistic approach to BMIs. In [31, 57, 50], probabilistic approaches are shown to be useful for some robust control analysis and synthesis problems even though these problems are not tractable in the deterministic framework. The randomized algorithms in these papers may be also useful for solving BMIs. In this case, we should exploit inherent properties of BMIs such as biconvexity.

Bibliography

[1] V. Balakrishnan and L. Vandenberghe, "Connections between duality in control theory and convex optimization," *Proc. 1995 Amer. Control Conf.*, pp. 4030–4034, 1995.

[2] V. Balakrishnan and S. Boyd, "Global optimization in control system analysis and synthesis," in C. T. Leondes, editor, *Control and Dynamic Systems: Advances in Theory and Applications*, pp. 1–55, Academic Press, 1992.

[3] V. Balakrishnan and S. Boyd and S. Balemi, "Branch and bound algorithm for computing the minimum stability degree of parameter-dependent linear systems," *Int. J. Robust and Nonlinear Control*, vol. 1, pp. 295–317, 1991.

[4] E. Beran, L. Vandenberghe and S. Boyd, "A global BMI algorithm based on the generalized Benders decomposition," *Proc. 1997 European Control Conf.*, 1997.

[5] Y. Boers, S. Weiland and A. A. H. Damen, "Constant feedback $\mathcal{H}_2$ control of systems with statistical Uncertainty," *Proc. 37th IEEE Conf. Decision and Control*, pp. 3328–3334, 1998.

[6] Y. Boers, S. Weiland and A. A. H. Damen, "Average $\mathcal{H}_2$ performance design by randomized algorithms," *Proc. 37th IEEE Conf. Decision and Control*, pp. 3347–3348, 1998.

[7] S. P. Boyd and C. H. Barratt, *Linear Controller Design: Limits of Performance*, Prentice Hall, 1991.

[8] S. Boyd, L. El Ghaoui, E. Feron and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*, SIAM, 1994.

[9] E. F. Camacho and C. Bordons, *Model Predictive Control in the Process Industry*, Springer-Verlag, 1995.

[10] D. Clarke, *Advances in Model-Based Predictive Control*, Oxford Univ. Press, 1994.

[11] A. V. Fiacco, *Introduction to Sensitivity and Stability Analysis in Nonlinear Programming*, Academic Press, 1983.

[12] C. A. Floudas and V. Visweswaran, "Primal-relaxed dual global optimization approach," *J. Optimization Theory and Applications*, vol. 78, no. 2, pp. 187–225, 1993.

[13] C. A. Floudas, *Nonlinear and Mixed-Integer Optimization: Fundamentals and Applications*, Oxford Univ. Press, 1995.

[14] J. H. Friedman, P. T. Kabamba and P. P. Khargonekar, "Worst-case and average $\mathcal{H}_2$ performance analysis against real constant parametric uncertainty," *Proc. 1994 Amer. Control Conf.*, pp. 2406–2410, 1994.

[15] H. Fujioka and T. Iwasaki, "BMI: a challenge to control synthesis based on nonconvex matrix inequalities," *J. SICE*, vol. 36, no. 11, pp. 762–767, 1997 (in Japanese).

[16] H. Fujioka and K. Hoshijima, "Bounds for the BMI eigenvalue problem – a good lower bound and a cheap upper bound –," *Trans. SICE*, vol. 33, no. 7, pp. 616–623, 1997.

[17] H. Fujioka, Y. Wakasa and Y. Yamamoto, "Optimization in control system design," *J. Operations Research Soc. Japan*, vol. 43, no. 1, pp. 48–70, 2000.

[18] M. Fukuda and M. Kojima, "Branch-and-cut algorithms for the bilinear matrix inequality eigenvalue problem," Research Rep. B-338, Dept. Mathematical and Computing Sciences, Tokyo Inst. of Tech., 1994.

[19] P. Gahinet, A. Nemirovski, A. J. Laub and M. Chilali, *LMI Control Toolbox: For Use with MATLAB*, The MathWorks Inc., 1995.

[20] A. M. Geoffrion, "Generalized Benders decomposition," *J. Optimization Theory and Applications*, vol. 10, no. 4, pp. 237–260, 1972.

[21] L. El Ghaoui and V. Balakrishnan, "Synthesis of fixed-structure controllers via numerical optimization," *Proc. 33rd IEEE Conf. Decision and Control*, pp. 2678–2683, 1994.

[22] C. E. Garcia, D. M. Prett and M. Morari, "Model predictive control: theory and practice – a survey," *Automatica*, vol. 25, no. 3, pp. 335–348, 1989.

[23] K. C. Goh, M. G. Safonov and G. P. Papavassilopoulos, "A global optimization approach for the BMI problem," *Proc. 33rd IEEE Conf. Decision and Control*, pp. 2009–2014, 1994.

[24] K. C. Goh, M. G. Safonov, L. Turan and M. G. Safonov, "μ/K_m synthesis via bilinear matrix inequalities," *Proc. 33rd IEEE Conf. Decision and Control*, pp. 2032–2037, 1994.

[25] K. C. Goh, L. Turan and M. G. Safonov, "Biaffine matrix inequality properties and computational methods," *Proc. 1994 Amer. Control Conf.*, pp. 850–855, 1994.

[26] K. M. Grigoriadis and R. E. Skelton, "Low-order control design for LMI problems using alternating projection methods," *Automatica*, vol. 32, no. 8, pp. 1117–1125, 1996.

[27] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge Univ. Press, 1991.

[28] R. Horst and H. Tuy, *Global Optimization: Deterministic Approaches*, Springer-Verlag, second, revised edition, 1993.

[29] M. Ikeda, G. Zhai and Y. Fujisaki, "Decentralized $\mathcal{H}_\infty$ controller design for large-scale-systems,' *Proc. 35th IEEE Conf. Decision and Control*, pp. 1–6, 1996.

[30] T. Iwasaki, *LMI and Control*, Shokodo, 1997 (in Japanese).

[31] P. P. Khargonekar and A. Tikku, "Probabilistic search algorithms for robust stability analysis and their complexity properties," in Y. Yamamoto and S. Hara, editors, *Learning, Control and Hybrid Systems*, pp. 25–45, Springer-Verlag, 1998.

[32] N. Karmarkar, "A new polynomial-time algorithm for linear programming," *Combinatorica*, vol. 4, no. 4, pp. 373–395, 1984.

[33] M. Kawanishi and T. Sugie, "BMI global optimization via BB method considering bilinearity," *Proc. 19th SICE symp. Dynamical System Theory*, pp. 59–62, 1996 (in Japanese).

[34] M. Kojima, "A primitive interior point algorithm for semidefinite programs in Mathematica," Research Rep. B-293, Dept. Mathematical and Computing Sciences, Tokyo Inst. of Tech., 1994.

[35] M. V. Kothare, V. Balakrishnan and M. Morari, "Robust constrained model predictive control using linear matrix inequalities," *Automatica*, vol. 32, no. 10, pp. 1361–1379, 1996.

[36] S. M. Liu and G. P. Papavassilopoulos, "Numerical experience with parallel algorithms for solving the BMI problem," *Proc. 13th World Congress of IFAC*, Vol. D, pp. 387–392, 1996.

[37] W. B. Liu and C. A. Floudas, "Generalized primal-relaxed dual approach for global optimization," *J. Optimization Theory and Applications*, vol. 90, no. 2, pp. 417–434, 1996.

[38] O. L. Mangasarian, *Nonlinear Programming*, McGraw-Hill, 1969.

[39] I. Masubuchi and A. Ohara, "LMI approach to control systems synthesis and semidefinite programming," *J. SICE*, vol. 35, no. 10, pp. 743–750, 1996.

[40] M. Mesbahi, G. P. Papavassilopoulos and M. G. Safonov, "Matrix cones, complementarity problems, and the bilinear matrix inequality," *Proc. 34th IEEE Conf. Decision and Control*, pp. 3102–3107, 1995.

[41] Yu. Nesterov and A. Nemirovsky, "A general approach to polynomial-time algorithms design for convex programming," Technical Rep., Centr. Econ. and Math. Inst., USSR Acad. Sci., 1988.

[42] Yu. Nesterov and A. Nemirovsky, *Interior-point Polynomial Algorithms in Convex Programming*, SIAM, 1994.

[43] G. Obinata, "Simultaneous optimal design of structural and control systems," *J. SICE*, vol. 36, no. 4, pp. 254–261, 1997 (in Japanese).

[44] A. Ohara and T. Sugie, "Control system design using convex optimization," *Systems, Control and Information*, vol. 36, no. 7, pp. 139–146, 1994 (in Japanese).

[45] M. Overton, "On minimizing the maximum eigenvalue of a symmetric matrix," *SIAM J. Matrix Analysis and Applications*, vol. 9, no. 2, pp. 256–268, 1988.

[46] S. J. Qin and T. A. Badgwell, "An overview of industrial model predictive control technology," *Proc. 5th Int. Conf. Chemical Process Control*, pp. 232–256, 1996.

[47] M. G. Safonov, K. C. Goh and J. H. Ly, "Control system synthesis via bilinear matrix inequalities," *Proc. 1994 Amer. Control Conf.*, pp. 45–49, 1994.

[48] C. Scherer, P. Gahinet and M. Chilali, "Multiobjective output-feedback control via LMI optimization," *IEEE Trans. Automatic Control*, vol. 42, no. 7, pp. 896–911, 1997.

[49] S. Takano, T. Watanabe and K. Yasuda, "Branch and bound technique for global solution of BMI," *Trans. SICE*, vol. 33, no. 7, pp. 701–708, 1997.

[50] R. Tempo, E. W. Bai and F. Dabbene, "Probabilistic robustness analysis: explicit bounds for the minimum number of samples," *Systems and Control Letters*, vol. 30, pp. 237–242, 1997.

[51] O. Toker and H. Özbay, "On the NP-hardness of solving bilinear matrix inequalities and simultaneous stabilization with static output feedback," *Proc. 1995 Amer. Control Conf.* pp. 2525–2526, 1995.

[52] K. Tsumura and S. Shin, "Robust control with distribution condition of system parameters," *Proc. 26th SICE Symp. Control Theory*, pp. 247–252, 1996.

[53] T. Tsuchiya, "New trends in optimization – interior point algorithms and related topics – III," *Systems, Control and Information*, vol. 42, no. 8, pp. 460–469, 1998.

[54] L. Vandenberghe and S. Balakrishnan, "Algorithms and software for LMI problems in control," *IEEE Control Systems*, vol. 17, no. 10, pp. 89–95, 1997.

[55] L. Vandenberghe and S. Boyd, "`SP`: software for semidefinite programming, user's guide, beta version," Available via anonymous ftp at `isl.stanford.edu` under `/pub/boyd/semidef_prog`, 1994.

[56] L. Vandenberghe and S. Boyd, "Semidefinite programming," *SIAM Review*, vol. 38, no. 1, pp. 49–95, 1996.

[57] M. Vidyasagar, "Randomized algorithms for robust controller synthesis using statistical learning theory," in Y. Yamamoto and S. Hara, editors, *Learning, Control and Hybrid Systems*, pp. 3–24, Springer-Verlag, 1998.

[58] Y. Wakasa, T. Tanino and T. Katayama, "Sensitivity analysis in semidefinite programming and its application to uncertain control systems," *Proc. 13th World Congress of IFAC*, vol. H, pp. 185–189, 1996.

[59] Y. Wakasa, T. Tanino and T. Katayama, "Sensitivity analysis in uncertain systems via semidefinite programming," *Trans. SICE*, vol. 32, no. 8, pp. 1212–1218, 1996 (in Japanese).

[60] Y. Wakasa, T. Tanino and T. Katayama, "Robust LMI-based model predictive control with slew rate constraints," *Proc. 2nd Asian Control Conf.*, vol. II, pp. 447–450, 1997.

[61] Y. Wakasa, "Global optimization methods for the BMI problem," *System, Control and Information*, vol. 43, no. 5, pp. 258–265, 1999 (in Japanese).

[62] Y. Wakasa, M. Sasaki, T. Tanino and T. Katayama, "A global approach to the BMI problem using the primal-relaxed dual method," *Trans. SICE*, vol. 35, no. 6, pp. 748–753, 1999 (in Japanese).

[63] Y. Wakasa and Y. Yamamoto, "Control system design considering a tradeoff between evaluated uncertainty ranges and control performance," *Asian J. Control*, vol. 1, no. 1, pp. 49–57, 1999.

[64] J. C. Willems, "Least squares stationary optimal control and the algebraic Riccati equation," *IEEE Trans. Automatic Control*, vol. 16, no. 6, pp. 621–634, 1971.

[65] Y. Yamada and S. Hara, "Global optimization for the matrix product eigenvalue problem," *Proc. 27th SICE symp. Control Theory*, pp. 343–348, 1998.